Group Theory

Groningen, 2nd year bachelor mathematics, 2016
(translation from original Dutch lecture notes, work in progress)

J. Top

Contents

preface

These lecture notes contain a translation into English of the Dutch lecture notes on Group Theory as they were used in the mathematics curriculum of Groningen University during the period 1993–2013. The original Dutch text may be found at `http://www.math.rug.nl/~top/alg1.pdf`.

Both the present text and the original are loosely based on another Dutch text on Group Theory, called *Algebra I*, written in the late 1970's at the university of Amsterdam by Prof.dr. F. Oort and Prof.dr. H.W. Lenstra.

Groningen, September 2016
Jaap Top

I THE INTEGERS

In this chapter the set of all integers $\mathbb{Z} = \{\ldots,-2,-1,0,1,2,3,\ldots\}$ is considered. Many elementary properties of these we learned in elementary and secondary education. However, in most cases formal proofs of such properties were not discussed. Such proofs form the main part of the present chapter, and may be seen as a repetition and extension of the same subject as it was treated during part of the first year bachelor's course 'Introduction to Mathematics'. In the second chapter of the present notes we will see how the developed theory about integers is used, for example, in order to obtain a better understanding of computations involving remainders upon division by a fixed integer. In turn, calculations using such remainders will be used to obtain criteria determining whether a given large integer is or is not a prime number.

Here as well as in subsequent chapters, many examples will be found illustrating how rather abstract definitions and proofs turn out to be quite applicable in concrete situations.

I.1 Division (with remainder)

In elementary school one encounters exercises like $100 : 7 = 14\,\mathrm{R}\,2$, meaning that 7 goes into 100 in total 14 times, leaving a remainder of 2. Behind problems of this sort is the following general fact.

I.1.1 Theorem. (Division with remainder.) *Let $a,b \in \mathbb{Z}$ with $b \neq 0$. Then $q,r \in \mathbb{Z}$ exist with*

$$a = qb + r \ \text{ and } \ 0 \leq r < |b|.$$

Moreover, these q,r are unique.

Proof. First we show existence of the required $q,r \in \mathbb{Z}$. Assume $a \geq 0$. We prove existence of q,r using mathematical induction with respect to a: in case $a = 0$ one may take $q = r = 0$. Now let $a > 0$ and use as induction hypothesis that $a - 1 \geq 0$ can be written as $a - 1 = \tilde{q}b + \tilde{r}$ with $0 \leq \tilde{r} < |b|$. Then $a = \tilde{q}b + \tilde{r} + 1$. Here evidently $0 \leq \tilde{r} + 1 \leq |b|$. In case $\tilde{r} + 1 < |b|$ we may take $q = \tilde{q}$ and $r = \tilde{r} + 1$. In the remaining case $\tilde{r} + 1 = |b|$ we have $a = \tilde{q}b + \tilde{r} + 1 = \tilde{q}b + |b| = (\tilde{q} + \frac{|b|}{b})b + 0$. Hence one can take $q = \tilde{q} + \frac{|b|}{b}$ and $r = 0$. Using the principle of mathematical induction this proves existence of q,r in the case $a \geq 0$.

If $a < 0$, then $-a > 0$, hence the argument above shows that there exist q',r' with $-a = q'b + r'$ and $0 \leq r' < |b|$. Then $a = (-q')b - r' = (-q' - \frac{|b|}{b})b + (|b| - r')$, so

we conclude that $q = -q'$ and $r = 0$ work in case $r' = 0$, while for $r' \neq 0$ one can take $q = -q' - \frac{|b|}{b}$ and $r = |b| - r'$.

It remains to prove uniqueness. Suppose $a = q_1 b + r_1 = q_2 b + r_2$ with $0 \leq r_1 \leq r_2 < |b|$. Then $0 \leq r_2 - r_1 \leq r_2 < |b|$, and also $r_2 - r_1 = b(q_1 - q_2)$. Hence $r_1 = r_2$, since otherwise $r_2 - r_1$ would be a positive multiple of $|b|$, contradicting $r_2 - r_1 < |b|$. As a consequence $b(q_1 - q_2) = 0$, and since $b \neq 0$ this implies $q_1 = q_2$. This proves the theorem. ∎

I.1.2 Remark. With a little knowledge about real numbers, a different argument may be given: partition the real line in intervals of length $|b|$, so

$$\mathbb{R} = \ldots \cup [-2|b|, -|b|) \cup [-|b|, 0) \cup [0, |b|) \cup \ldots\ldots$$

Then $a \in \mathbb{R}$ is in exactly one such interval, hence can be written as $a = qb + r$ with $0 \leq r < |b|$.

A reason to prefer the proof using induction over the argument involving real numbers, is that conceptually $\mathbb{R}$ is much more difficult than $\mathbb{Z}$. In fact, one can *construct* $\mathbb{R}$ by first constructing the rational numbers $\mathbb{Q}$ starting from $\mathbb{Z}$, and then building $\mathbb{R}$ from $\mathbb{Q}$ via a technique called 'completion'.

I.1.3 Definition. Let $a, b \in \mathbb{Z}$. We say that a *divides* the integer b, if $q \in \mathbb{Z}$ exists such that $b = qa$. This is denoted as $a|b$. In case no such q exists, we write $a \not| b$ (and we say: a does *not* divide b).

Instead of a divides b one also says that a is a *divisor* of b, or that a is a *factor* of b, or that b is a *multiple* of a, or that b is *divisible* by a. For example $17|-153$ and $0|0$ and $-2 \not| 101$ and $0 \not| 3$.

We present some elementary properties of divisibility.

I.1.4 Proposition. *For $a, b, c \in \mathbb{Z}$ one has:*

1. *If $a|b$ and $b|c$ then also $a|c$.*
2. *If $a|b$ and $a|c$ then also $a|b \pm c$.*
3. *$a|0$ and $1|a$.*
4. *$0|a$ if and only if $a = 0$.*
5. *If for $b \neq 0$ it holds that $a|b$, then $|a| \leq |b|$.*

Proof. These properties are immediate consequences of the definition. As an example, $a|b$ and $a|c$ implies that $p, q \in \mathbb{Z}$ exist with $b = pa$ and $c = qa$, and hence it follows that $b \pm c = pa \pm qa = (p \pm q)a$, so $a|b \pm c$. ∎

A consequence of the last of the properties mentioned in Proposition I.1.4, is that an integer $a \neq 0$ has only finitely many divisors; the largest of these is evidently $|a|$. Given another integer, say b, then in particular a and b have only finitely many divisors in common (two of the common divisors are of course 1 and -1). If one considers common multiples of two integers a, b, then in case a or b equals 0, the fourth property mentioned in Proposition I.1.4 implies that 0 is the only common multiple. However, in case $ab \neq 0$ the integers a and b have *positive* common multiples, for example $|ab|$. This leads to the following definition:

I.1.5 Definition. Let $a, b \in \mathbb{Z}$. In case a and b are not both equal to 0, the *greatest common divisor* of a and b is defined as the largest integer that is a divisor of both a and b. This integer is denoted as $\gcd(a, b)$. Furthermore, by definition we say $\gcd(0, 0) = 0$.

The *least common multiple*, notation $\mathrm{lcm}(a, b)$ of a and b is by definition 0 in case $ab = 0$, and it is the smallest positive integer k with the property $a|k$ and $b|k$ if both a and b are nonzero.

Two integers a, b we call *coprime* or also *relative prime* if $\gcd(a, b) = 1$.

I.1.6 Example. One has $\gcd(a,b) = \gcd(b,a)$ and $\gcd(a,0) = |a|$. Since a and $|a|$ have the same divisors, $\gcd(a,b) = \gcd(|a|,b) = \gcd(a,|b|) = \gcd(|a|,|b|)$. It is in general a difficult task to compute directly from the definition a greatest common divisor. Try, for example, to check that $\gcd(35581,46189) = 221$.

One easily constructs similar examples with the lcm; for instance, one encounters them in elementary school when one tries to find a common denominator for two fractions. In the remainder of this section we will only discuss the gcd; in Section I.2 we return to the notion lcm once more. ∎

It turns out that a surprisingly simple and efficient algorithm exists for computing $\gcd(a,b)$. This dates back from the Greek mathematician Euclid who lived around 300 B.C. The algorithm runs as follows.

I.1.7 Theorem. (The Euclidean algorithm.) *The following algorithm finds in finitely many steps the greatest common divisor of two integers a,b:*

```
gcd:=proc(a::integer,b::integer)::integer;
        local rn,ro,help;
        ro:=max(abs(a),abs(b)); rn:=min(abs(a),abs(b));
        while rn<>0 do
            do help:=ro; ro:=rn; rn:=help mod rn end do;
        return ro
    end proc;
```

Proof. To understand this program, we check what happens during the 'while-loop'. Each time this loop is executed, the pair of integers (ro, rn) is replaced by (rn, ro mod rn). Here ro mod rn is the remainder upon dividing ro by rn. (Note, by the way: in many programming languages, for *negative* a the result of a mod b is *not* the remainder r as given in Theorem I.1.1, but it is $r-|b|$. The code above is written in Maple; here this problem does not occur.) In particular, at the start of the 'loop' it holds that ro, rn ≥ 0, and each time the loop is executed, rn becomes strictly smaller. Hence the computer program terminates. To show that the program indeed computes the greatest common divisor of a and b, we show that moreover when entering the 'loop', each time it holds that $\gcd(\text{ro},\text{rn}) = \gcd(a,b)$. This will be done in Lemma I.1.8 below. Accepting the lemma, it follows that after the last execution of the 'loop' rn $= 0$ and $\gcd(a,b) = \gcd(\text{ro},\text{rn}) = \gcd(\text{ro},0) = \text{ro}$. In other words, indeed the algorithm outputs the greatest common divisor of a and b. ∎

I.1.8 Lemma. *For $a,b,q,r \in \mathbb{Z}$ with $a = qb + r$ one has $\gcd(a,b) = \gcd(b,r)$.*

Proof. We will show that the set of common divisors of a and b equals the set of common divisors of b and r. The definition of greatest common divisor then implies the lemma.

If $d|a$ and $d|b$, then also $d|a-qb = r$. Hence common divisors of a and b are also common divisors of b and r.

Vice versa, if $d|b$ and $d|r$, then also $d|qb+r = a$. Hence the common divisors of b and r are also common divisors of a and b. This proves the lemma. ∎

I.1.9 Example.

$$\begin{array}{lll} \gcd(1057,315) & = \gcd(3\cdot 315+112,315) & = \\ & = \gcd(315,112) & = \gcd(2\cdot 112+91,112) = \\ & = \gcd(112,91) & = \gcd(91+21,91) = \\ & = \gcd(91,21) & = \gcd(4\cdot 21+7,21) = \\ & = \gcd(21,7) & = \gcd(7,0) = 7. \end{array}$$

∎

We now discuss the efficiency of the Euclidean algorithm.

I.1.10 Theorem. (G. Lamé, 1844, French mathematician.) *If $a > b > 0$, then the number of divisions with remainder performed by the Euclidean algorithm when determining* $\gcd(a,b)$ *is at most* 5 *times the number of decimal digits of b.*

Proof. Write $r_0 = a$ and $r_1 = b$. The algorithm computes one by one

$$\begin{array}{lcll} r_0 & = & q_0 r_1 + r_2 & (0 < r_2 < r_1) \\ r_1 & = & q_1 r_2 + r_3 & (0 < r_3 < r_2) \\ r_2 & = & q_2 r_3 + r_4 & (0 < r_4 < r_3) \\ & & \vdots & \\ r_{n-2} & = & q_{n-2} r_{n-1} + r_n & (0 < r_n < r_{n-1}) \\ r_{n-1} & = & q_{n-1} r_n + 0. & \end{array}$$

The number of divisions with remainder is therefore exactly n.

To estimate n we use the so-called Fibonacci sequence $(f_i)_{i\geq 0}$, defined inductively by $f_0 = f_1 = 1$ and $f_{i+2} = f_{i+1} + f_i$ for $i \geq 0$. So this is the sequence $1,1,2,3,5,8,13,21,34,\ldots\ldots$.

Using mathematical induction, we now show $r_{n-i} \geq f_{i+1}$ for $i = 0,\ldots,n-1$. The case $i = 0$ is trivial. The case $i = 1$: since $0 < r_n < r_{n-1}$, it follows that $q_{n-1} > 1$ hence $r_{n-1} \geq 2r_n \geq 2 = f_2$. Now assume $i > 1$ and use the induction hypothesis for $i-1$ and $i-2$. It follows that $r_{n-i} = q_{n-i} r_{n-(i-1)} + r_{n-(i-2)} \geq r_{n-(i-1)} + r_{n-(i-2)} \geq f_i + f_{i-1} = f_{i+1}$. This completes the induction.

In particular, this shows that $b = r_1 \geq f_n$.

To complete the proof of the theorem, we again use induction to show $f_{5i+1} > 10^i$ for $i \geq 1$. For $i = 1$ this is correct, since $f_6 = 13 > 10$. Assuming the inequality for $i \geq 1$, it follows that

$$\begin{array}{ll} f_{5(i+1)+1} & = f_{5i+6} = f_{5i+5} + f_{5i+4} = f_{5i+4} + 2f_{5i+3} + f_{5i+2} = f_{5i+3} + 3f_{5i+2} + 3f_{5i+1} + f_{5i} \\ & = f_{5i+2} + 7f_{5i+1} + 4f_{5i} = 8f_{5i+1} + 5f_{5i} > 8f_{5i+1} + 2f_{5i} + 2f_{5i-1} = 10f_{5i+1} \\ & > 10 \cdot 10^i = 10^{i+1}. \end{array}$$

Now write $n = 5m + k$ with $1 \leq k \leq 5$. Then $b \geq f_n \geq f_{5m+1} > 10^m$. This shows that the number of decimal digits of b is at least $m + 1 \geq n/5$, so n is at most 5 times the number of decimal digits of b, which is what we wanted to prove. ∎

An application of the Euclidean algorithm which will also be quite useful later on in 'modular arithmetic', is that one can construct solutions in integers of certain linear equations with it:

I.1.11 Theorem. (Bachet-Bézout; named after the French mathematicians Claude Gaspard Bachet 1581–1638 and Étienne Bézout 1730–1783.) *For $a,b \in \mathbb{Z}$ there exist integers $x,y \in \mathbb{Z}$ such that $ax + by = \gcd(a,b)$.*

Proof. In case $ab = 0$ this is immediate from the definition of the gcd. Now assume $a \neq 0 \neq b$. Write $r_0 = |a|$ en $r_1 = |b|$ and let n be the number of divisions with remainder computed during the execution of the Euclidean algorithm. Then we have a sequence

$$\begin{array}{lcll} r_0 & = & q_0 r_1 + r_2 & (0 < r_2 < r_1) \\ r_1 & = & q_1 r_2 + r_3 & (0 < r_3 < r_2) \\ r_2 & = & q_2 r_3 + r_4 & (0 < r_4 < r_3) \\ & & \vdots & \\ r_{n-2} & = & q_{n-2} r_{n-1} + r_n & (0 < r_n < r_{n-1}) \\ r_{n-1} & = & q_{n-1} r_n + 0 & \end{array}$$

as used in the previous proof. The $n-1$st equality here presents $r_n = \gcd(a,b)$ as a linear combination of r_{n-1} and r_{n-2} with integer coefficients. Using the $n-2$nd equality one can write r_{n-1} as a combination of r_{n-2} and r_{n-3}, so this results in a presentation of $\gcd(a,b)$ as an integer linear combination of r_{n-2} and r_{n-3}. Working upward, one consecutively eliminates $r_{n-2}, r_{n-3}, \ldots, r_3, r_2$. What remains is a relation $\gcd(a,b) = xr_1 + yr_0$. By changing the sign of x and/or y if necessary, this yields an equality $ax + by = \gcd(a,b)$ as desired. ∎

The argument presented above is completely constructive. The next algorithm finds at the same time the greatest common divisor of $a, b \in \mathbb{Z}$, and integers x, y such that $ax + by = \gcd(a,b)$. In fact the algorithm does not consider the sequence $r_n, r_{n-1}, \ldots, r_0$ as discussed in the proof, but rather the sequence in reversed order $r_0, r_1, r_2, \ldots, r_n$. Namely, with every new r_i immediately $x_i, y_i \in \mathbb{Z}$ are found such that $x_i a + y_i b = r_i$. In the nth step, these x_n, y_n are the required x, y. Check for yourself that indeed the algorithm works.

```
# Here we find gcd(a,b), and write it as xa+yb.
if a=0
then x:=0; y:=1; gcd:=abs(b)
else if b=0
    then x:=1; y:=0; gcd:=abs(a)
    else # a and b are both nonzero in this case
        ro:=abs(a); xo:=sign(a); yo:=0;
        rn:=abs(b); x:=0; y:=sign(b);
        while rn<>0
            do
               q:=floor(ro/rn); help:=rn; rn:=ro-q*rn; ro:=help;
               help:=x; x:=xo-q*x; xo:=help;
               help:=y; y:=yo-q*y; yo:=help
            end do;
        gcd:=ro; x:=xo; y:=yo
    end if
end if; print(gcd,x,y);
```

I.1.12 Example. In Example I.1.9 we saw that $\gcd(1057, 315) = 7$. We now construct integers x, y such that $1057x + 315y = 7$ using the algorithm above. To this end, consider the following equalities:

$$\begin{array}{rcrcrl}
1\cdot 1057 & + & 0\cdot 315 & = & 1057 & \\
0\cdot 1057 & + & 1\cdot 315 & = & 315 & \text{(subtract this 3 times from the previous:)} \\
1\cdot 1057 & + & -3\cdot 315 & = & 112 & \text{(this one 2 times from the previous:)} \\
-2\cdot 1057 & + & 7\cdot 315 & = & 91 & \text{(this once from the previous:)} \\
3\cdot 1057 & + & -10\cdot 315 & = & 21 & \text{(this 4 times from the previous:)} \\
-14\cdot 1057 & + & 47\cdot 315 & = & 7. &
\end{array}$$

The solution found here is by far not the only one. Is for example also (x', y') a solution, then with respect to the standard inner product on $\mathbb{R}^2$ one has

$$\langle \begin{pmatrix} 1057 \\ 315 \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \rangle = 7 = \langle \begin{pmatrix} 1057 \\ 315 \end{pmatrix}, \begin{pmatrix} x' \\ y' \end{pmatrix} \rangle,$$

so the vector $\binom{x-x'}{y-y'}$ is perpendicular to $\binom{1057}{315}$. Using this it is not hard to find all solutions in integers of the equation $1057x + 315y = 7$. ∎

Using the existence of $x, y \in \mathbb{Z}$ with $ax + by = \gcd(a,b)$, it is easy to derive some further results:

I.1.13 Corollary. *Let $a, b \in \mathbb{Z}$ and put $d = \gcd(a, b)$. Every integer that is a divisor of a as well as of b is also a divisor of d. Vice versa, any divisor of d is a common divisor of a and b.*

Proof. Since d divides both a and b, the first property in Proposition I.1.4 shows that any divisor of d divides a and b as well.

For the other assertion, write $d = ax + by$ for certain $x, y \in \mathbb{Z}$. If $c|a$ and $c|b$ then also $c|ax + by = d$. This proves the corollary. ∎

I.1.14 Corollary. *Let $a, b \in \mathbb{Z}$. Then a, b are coprime if and only if $x, y \in \mathbb{Z}$ exist with $ax + by = 1$.*

Proof. If a, b are coprime, then by definition $\gcd(a, b) = 1$. So Theorem I.1.11 implies the existence of $x, y \in \mathbb{Z}$ with $ax + by = 1$.

Vice versa, if $ax + by = 1$ for some $x, y \in \mathbb{Z}$, put $d = \gcd(a, b)$. Then $d \geq 0$ and $d|a$ and $d|b$, hence $d|ax + by = 1$, so $d = 1$. ∎

I.1.15 Corollary. *For $a, b, c \in \mathbb{Z}$ with $\gcd(a, b) = 1$ the following holds: if $a|bc$, then $a|c$.*

Proof. Take $x, y \in \mathbb{Z}$ with $ax + by = 1$. If $a|bc$, then also $a|axc + byc = (ax + by)c = c$. ∎

I.2 Prime factorization

I.2.1 Definition. A *prime number* (prime) is an integer larger than 1, that has only 1 and itself as positive divisors.

I.2.2 Example. Small prime numbers such as $2, 3, 5, \ldots$ are well known. For much larger integers it is in general not easy to check whether they are prime. We know for example that $2^{524288} + 1$ is *not* prime, but $2^{859433} - 1$ is, and $2^{1048576} + 1$ is not. These numbers consist of, respectively 157827, 258716, and 315653 decimal digits. The exponents here are 2^{19}, the prime 859433, and 2^{20}. In 1962 the Swedish mathematician Hans Riesel found a divisor of the first mentioned number, namely $33629 \cdot 2^{21} + 1$. In January 1994 the second mentioned number was proven to be prime by the Americans David Slowinski and Paul Gage. At the moment (August 2016) we know much larger prime numbers, such as $2^{74207281} - 1$, consisting of 22338618 decimal digits. This immense example was discovered in January 2016 by the American mathematician Curtis Cooper. Concerning the number $2^{2^{20}} + 1$ given above, in 1987 the American mathematicians Jeff Young and Duncan Buell showed it is not a prime. However, until now (August 2016) nobody was able to find a nontrivial factor. The smallest integer of the form $1 + 2^m$ for which at present it is unknown whether it is prime, is the one with $m = 2^{33}$. ∎

I.2.3 Theorem. *If p is prime, and $a, b \in \mathbb{Z}$ such that $p|ab$, one has $p|a$ or $p|b$.*

Proof. Put $d = \gcd(a, p)$. Since $p \neq 0$, this d is positive. Moreover d divides p. The definition of a prime therefore implies $d = 1$ or $d = p$. In the first case Corollary I.1.15 shows $p|b$. In the second case one has $p = d|a$. ∎

I.2.4 Corollary. *If p is prime, and $a_1, \ldots, a_n$ are integers such that $p|a_1 a_2 \cdot \ldots \cdot a_n$, then an index k exists with $1 \leq k \leq n$ such that $p|a_k$.*

Proof. This can be shown by induction with respect to n. For $n = 1$ the statement is obvious (and the case $n = 2$ is shown in Theorem I.2.3). Now take $n \geq 3$ and assume the statement for products of $< n$ factors. If $p|a_1a_2 \cdot \ldots \cdot a_n = (a_1) \cdot (a_2 \cdot \ldots \cdot a_n)$, Theorem I.2.3 implies that $p|a_1$ or $p|a_2 \cdot \ldots \cdot a_n$. In the first case we are done, and in the second case we use the induction hypothesis. ∎

Aided by the above properties of primes, we now show a result called the 'main theorem of arithmetic':

I.2.5 Theorem. (unique prime factorisation) *Every integer greater than* 1 *can be written as a product of primes. This product is unique up to the order of the factors.*

Proof. We first show that any $n \in \mathbb{Z}$ with $n > 1$ can be written as a product of primes. We use induction w.r.t. n: the case $n = 2$ is clear. Let $n > 2$ and suppose every integer greater than 1 and smaller than n can be written as a product of primes. If n is a prime number, we are done. If n not prime, then $n = n_1n_2$ and $1 < n_1, n_2 < n$. The induction hypothesis implies that both n_1 and n_2 are products of primes, so n is as well.

Next we show uniqueness. Suppose uniqueness does not hold, and take n the smallest integer > 1 allowing more than one factorisation into primes, say

$$n = p_1p_2 \cdot \ldots \cdot p_t = q_1q_2 \cdot \ldots \cdot q_s,$$

with primes p_i, q_j. Then $p_1|n = q_1q_2 \cdot \ldots \cdot q_s$. Since n allows more than one factorisation, n is not prime, hence $s, t > 1$ so in particular $n/p_1 > 1$. Corollary I.2.4 implies $p_1|q_k$ for some k. Since q_k is prime, we conclude that $p_1 = q_k$. Dividing the given factorisations by their common factor $p_1 = q_k$, it follows that n/p_1 allows two factorisations as well. This contradicts the minimality of n. Hence the theorem is proven. ∎

Although we saw in Example I.2.2 that finding large primes is not easy, yet the following result is very old.

I.2.6 Theorem. (Euclid) *There exist infinitely many primes.*

Proof. Suppose we have $n \geq 1$ pairwise different primes $p_1, \ldots, p_n$. Consider the integer $N = (p_1 \cdot p_2 \cdot \ldots \cdot p_n) + 1$. Take a prime q in the prime factorisation of N. Then q differs from each p_i, since otherwise q divides both N and $N - 1 = p_1 \cdot \ldots \cdot p_n$ hence also their difference $N - (N - 1) = 1$ which is impossible. We therefore conclude from the existence of n primes that also $n + 1$ primes exist. The result follows. ∎

I.2.7 Remark. It is *not* true that N as constructed in the proof of Theorem I.2.6 is necessarily itself a prime. For example, given a finite set of *odd* primes, the product plus 1 is even (and > 2). Also, starting from the set of primes {2}, then repeatedly taking 1 plus the product of the set of primes yields $2, 3, 7, 43, 1807 = 13 \cdot 139$.

I.2.8 Definition. If p is prime and $a \in \mathbb{Z}$ not equal to 0 or ± 1, then we write $v_p(a)$ for the number of times p appears in the prime factorisation of $|a|$. Moreover, we put $v_p(1) = v_p(-1) = 0$ and $v_p(0) = \infty$.

The number $v_p(a)$ is usually called the *valuation* of a at p. Since a prime factorisation is unique up to the order of the primes, $v_p(a)$ is well defined. If $a \in \mathbb{Z}$ is not zero, then the definition implies $|a| = \prod p^{v_p(a)}$. Here the product is taken over *all* primes p, and although according to Theorem I.2.6 there is an infinitude of primes, yet the product is well defined. Namely, only finitely many primes occur in the prime factorisation of $|a|$. For all other primes p one has $v_p(a) = 0$ hence $p^{v_p(a)} = 1$.

I.2.9 Corollary. *Let a, b be integers.*

1. *For every prime p we have $v_p(ab) = v_p(a) + v_p(b)$.*
2. *$a|b$ if and only if every prime p satisfies $v_p(a) \le v_p(b)$.*
3. *If a and b are not both zero, then*
$$\gcd(a,b) = \prod_{p \text{ prime}} p^{\min\{v_p(a), v_p(b)\}}.$$
4. *If $a \neq 0$ and also $b \neq 0$, then*
$$\operatorname{lcm}(a,b) = \prod_{p \text{ prime}} p^{\max\{v_p(a), v_p(b)\}}.$$
5. *If $a|c$ and $b|c$, then $\operatorname{lcm}(a,b)|c$.*
6. *$\gcd(a,b) \cdot \operatorname{lcm}(a,b) = |ab|$.*

Proof. 1: The statement is true if a and/or b equals zero, since infinity plus infinity and also infinity plus any finite number equal infinity. The remaining case follows from the equality
$$|ab| = |a| \cdot |b| = \prod_p p^{v_p(a)} \cdot \prod_p p^{v_p(b)} = \prod_p p^{v_p(a)+v_p(b)}.$$
2: If $a|b$, then $b = qa$ for some $q \in \mathbb{Z}$, and the first part of the corollary then implies $v_p(a) \le v_p(a) + v_p(q) = v_p(qa) = v_p(b)$, for all primes p.

Vice versa, assume $v_p(a) \le v_p(b)$ for all primes p . Certainly $a|b$ if $b = 0$. In case $a = 0$ the assumption implies $v_p(b) = \infty$, hence $b = 0$, so again $a|b$. If a and b are both nonzero, then $q = \prod_p p^{v_p(b)-v_p(a)}$ is a well defined integer. By definition $|b|$ and $|qa|$ have the same prime factorisation, hence $b = \pm qa$, which means $a|b$.
3: Using the result above, integers d which divide both a and b are precisely all integers with the properties $v_p(d) \le v_p(a)$ and $v_p(d) \le v_p(b)$ for all primes p. Since either $a \neq 0$ or $b \neq 0$ (or both), $\min\{v_p(a), v_p(b)\}$ is finite for all primes p, and nonzero for only finitely many primes p. So $\prod_p p^{\min\{v_p(a),v_p(b)\}}$ is a well defined integer, dividing both a and b and moreover at least as large as any other common divisor. Hence it equals $\gcd(a,b)$.
4: a and b are both nonzero, hence $v_p(a)$ and $v_p(b)$ are both finite for all primes p and nonzero for only finitely many primes p. Therefore $k = \prod_p p^{\max\{v_p(a),v_p(b)\}}$ is well defined and positive. By the second property above, k is a multiple of both a and b. Every common multiple c satisfies by the same property $v_p(c) \ge v_p(a)$ and $v_p(c) \ge v_p(b)$ for all primes p, hence k is the smallest positive common multiple, i.e., $k = \operatorname{lcm}(a,b)$.
5: If $ab = 0$ and $a|c$ and $b|c$, then $c = 0$ so in this case the assertion holds. If $ab \neq 0$, then the assertion follows by combining the 2nd and 4th property shown above.
6: This holds in case $ab = 0$ since than by definition $\operatorname{lcm}(a,b) = 0$. For $ab \neq 0$, one has
$$\begin{aligned} |ab| &= \textstyle\prod_p p^{v_p(a)} \cdot \prod_p p^{v_p(b)} = \prod_p p^{v_p(a)+v_p(b)} \\ &= \textstyle\prod_p p^{\min\{v_p(a),v_p(b)\}+\max\{v_p(a),v_p(b)\}} \end{aligned}$$
which using 3) and 4) equals $\gcd(a,b) \cdot \operatorname{lcm}(a,b)$. ∎

I.2.10 Remark. The formulas given in Corollary I.2.9 provide a method to compute $\gcd(a,b)$ and $\operatorname{lcm}(a,b)$ provided the prime factorisation of a and b is known. In general this is of course not the case, and is in many instances much more difficult than directly computing $\gcd(a,b)$ with the Euclidean algorithm. The knowledge of $\gcd(a,b)$ allows us to find $\operatorname{lcm}(a,b)$ by the formula $\operatorname{lcm}(a,b) = |ab|/\gcd(a,b)$. Observe that to *prove* this formula, unique prime factorisation was used. However to compute $\operatorname{lcm}(a,b)$ by means of the formula, this factorisation is no longer needed.

I.3 Exercises

1. ('The b-ary system'). Let $a, b \in \mathbb{Z}$ with $a \geq 1$ and $b \geq 2$. Show that $t \in \mathbb{Z}$ exists with $t \geq 0$, and moreover integers $c_0, c_1, \ldots, c_t \in \{0, 1, \ldots, b-1\}$ such that $c_t \neq 0$ and
$$a = c_t b^t + \ldots + c_2 b^2 + c_1 b + c_0.$$
Show that these $t, c_0, \ldots, c_t$ are *unique*.
2. Prove that in case $a, b \in \mathbb{Z}$ are not both equal to zero, then $a/\gcd(a,b)$ and $b/\gcd(a,b)$ are coprime.
3. Determine $d = \gcd(3354, 3081)$ and find $x, y \in \mathbb{Z}$ with $3354x + 3081y = d$. Next, find *all* solutions in integers of this equation.
4. Given is the Fibonacci sequence $(f_n)_{n \geq 0}$ defined by $f_0 = f_1 = 1$ and $f_{n+2} = f_{n+1} + f_n$ for $n \geq 0$. How many divisions with remainder are performed by the Euclidean algorithm when computing $\gcd(f_n, f_{n+1})$? Show that $\gcd(f_n, f_{n+1}) = 1$ for all $n \geq 0$.
5. Let $n \in \mathbb{Z}$, $n \geq 2$. Show that n is prime if and only if n has no divisor d such that $1 < d \leq \sqrt{n}$.
6. Prove that infinitely many primes exist which leave a remainder 3 upon division by 4.
7. Prove that infinitely many primes p exist with the property that $p - 2$ is not prime.
8. Prove that for $a, b, c \in \mathbb{Z}$ one has:
   (a) If $\gcd(a,b) = \gcd(a,c) = 1$, then $\gcd(a, bc) = 1$.
   (b) If a and b are coprime and both divide c, then their product divides c.
   (c) Is $c \geq 0$, then $\gcd(ac, bc) = c \cdot \gcd(a,b)$.
9. Take $a, b \in \mathbb{Z}$ both positive.
   (a) Let r be the remainder upon dividing a by b. Show that $2^r - 1$ is the remainder upon dividing $2^a - 1$ by $2^b - 1$.
   (b) Show that $2^b - 1 | 2^a - 1$ if and only if $b | a$.
   (c) Prove that $\gcd(2^a - 1, 2^b - 1) = 2^{\gcd(a,b)} - 1$.
   (d) Are the above assertions still true when one replaces '2' by some integer $c > 2$?
10. Given are $a, b, n \in \mathbb{Z}$ with $n \geq 0$.
    (a) Show that $a - b | a^n - b^n$.
    (b) Show that if n is odd, then $a + b | a^n + b^n$.
    (c) Now take $b = 1$ and assume $a > 1$ and $n > 1$. Prove that if $a^n - 1$ is prime, then $a = 2$ and n is prime.
    (d) Take $a = 2$ and $n = 11$ and verify that $a^n - 1$ is not prime. So apparently the converse of the property above does not necessarily hold.
11. Show (using part (b) of the previous exercise) that if $2^n + 1$ is prime, then $n = 2^k$ for an integer $k \geq 0$.
12. Prove that if $n^4 + 4^n$ is prime for some $n \in \mathbb{Z}, n \geq 1$, then $n = 1$.
13. (a) Show that any prime $p > 3$ satisfies $24 | p^2 - 1$.
    (b) Show that if p_1, p_2, p_3, p_4, p_5 are (not necessarily distinct) primes, and $p_1 p_2 p_3 p_4 p_5 + 1 = p^2$ for some prime p, then $p = 7$ or $p = 11$ or $p = 13$.
14. Suppose that $p_1, \ldots, p_n$ are pairwise distinct primes. Show that the real numbers $\log(p_1), \ldots, \log(p_n)$ are linearly independent over $\mathbb{Q}$, i.e., no $a_1, \ldots, a_n \in \mathbb{Q}$ exist which are not all zero, while $a_1 \log(p_1) + \ldots + a_n \log(p_n) = 0$.

II MODULAR ARITHMETIC

In this chapter we develop the basic properties of calculations with remainders upon division.

II.1 Residue classes modulo N

Let N be an arbitrary positive integer.

II.1.1 Definition. Two integers a, b are called *congruent modulo* N if $N|a-b$. This is denoted by $a \equiv b \bmod N$.

Integers a, b are congruent modulo N precisely when they have the same remainder upon division by N. Indeed, if $a = q_1N + r_1$ and $b = q_2N + r_2$ for certain $0 \le r_1, r_2 < N$, then the statement $N|a-b$ is equivalent to $N|r_1 - r_2$. Since $r_1 - r_2$ is strictly between $-N$ and $+N$, we have $N|r_1 - r_2$ if and only if $r_1 - r_2 = 0$, i.e., $r_1 = r_2$.

The relation 'being congruent modulo N' is an equivalence relation on $\mathbb{Z}$. Indeed, if a and b have the same remainder upon division by N, then the same holds for b and a, hence the relation is symmetric. The relation is reflexive, which means here that a and a have the same remainder upon division by N. Finally, if a and b yield the same remainder, and so do b and c, then clearly the same holds for a and c showing that the relation is transitive.

Any equivalence relation partitions a set into a union of pairwise disjoint subsets. In our case these subsets are called residue classes modulo N. Explicitly:

II.1.2 Definition. For $a \in \mathbb{Z}$ the *residue class of* a *modulo* N is given as

$$\{b \in \mathbb{Z} \mid b \equiv a \bmod N\}.$$

We denote this residue class by $a \bmod N$ or also, if no confusion is likely concerning the N in question, by $\overline{a}$.

So, by definition $a \bmod N$ is a *subset* of $\mathbb{Z}$. If $a = qN + r$, then $a \equiv r \bmod N$, and the residue class $r \bmod N$ equals $a \bmod N$. Hence there are equally many distinct residue classes modulo N as there are possible remainders upon division by N, namely N such classes. The residue class of a modulo N consists of all integers of the form $a + Nk$ for some $k \in \mathbb{Z}$, hence we can also write

$$a \bmod N = \overline{a} = a + N\mathbb{Z}.$$

II.1.3 Example. For $N = 4$ as explained above, there are 4 distinct residue classes, namely 0 mod 4 and 1 mod 4 and 2 mod 4 and 3 mod 4. As discussed, these are 4 pairwise disjoint subsets of $\mathbb{Z}$ whose union is all of $\mathbb{Z}$. They are:

$$\begin{aligned}
&\{ \ \ldots\ldots, -284, \ldots, -8, -4, 0, 4, \ldots, 1016, \ldots\ldots \ \} = 0 \bmod 4,\\
&\{ \ \ldots\ldots, -283, \ldots, -7, -3, 1, \ldots, 1017, \ldots\ldots \ \} = 1 \bmod 4,\\
&\{ \ \ldots\ldots, -282, \ldots, -6, -2, 2, \ldots, 1018, \ldots\ldots \ \} = 2 \bmod 4,\\
&\{ \ \ldots\ldots, -281, \ldots, -5, -1, 3, \ldots, 1019, \ldots\ldots \ \} = 3 \bmod 4.
\end{aligned}$$

The residue class 17 mod 4 equals 1 mod 4. Using the notation $\overline{a} = a \bmod 4$ this is expressed as $\overline{17} = \overline{1}$. Similarly one has $\overline{-1001} = \overline{3}$. ▬■

The following elementary property is an immediate consequence of generalities concerning equivalence relations. Nevertheless we present a proof, illustrating how the given definitions are used.

II.1.4 Lemma. *For $a, b \in \mathbb{Z}$ one has $a \bmod N = b \bmod N$ if and only if $a \equiv b \bmod N$.*

Proof. Assume $a \bmod N = b \bmod N$. Since a is an element of the residue class $a \bmod N = b \bmod N$, by definition $a \equiv b \bmod N$.

Vice versa, suppose $a \equiv b \bmod N$. As we saw, this means that a and b yield the same remainder upon division by N. Hence for any $c \in \mathbb{Z}$ it holds that c is in the residue class $a \bmod N$ if and only if c and a yield the same remainder upon division by N, which is equivalent to c leaving the same remainder as b, therefore to c being in $b \bmod N$. This shows $a \bmod N = b \bmod N$. ■

II.1.5 Theorem. *Let $\overline{a_1}, \overline{a_2}, \overline{b_1}, \overline{b_2}$ be residue classes modulo N for $a_1, a_2, b_1, b_2 \in \mathbb{Z}$, and $\overline{a_1} = \overline{a_2}$ and $\overline{b_1} = \overline{b_2}$. Then $\overline{a_1 + b_1} = \overline{a_2 + b_2}$ and $\overline{a_1 b_1} = \overline{a_2 b_2}$.*

Proof. From $\overline{a_1} = \overline{a_2}$ and $\overline{b_1} = \overline{b_2}$ it follows by Lemma II.1.4 that $q, q' \in \mathbb{Z}$ exist with $a_2 = a_1 + Nq$ and $b_2 = b_1 + Nq'$. Hence $a_2 + b_2 = a_1 + b_1 + N(q + q')$, which by Lemma II.1.4 implies $\overline{a_1 + b_1} = \overline{a_2 + b_2}$. Furthermore,

$$a_2 b_2 = (a_1 + Nq)(a_2 + Nq') = a_1 a_2 + N(a_1 q' + q a_2 + Nqq'),$$

and the same reasoning implies $\overline{a_1 b_1} = \overline{a_2 b_2}$. ■

II.1.6 Definition. (*Adding and multiplying residue classes.*) The set of residue classes modulo N is denoted by $\mathbb{Z}/N\mathbb{Z}$. For $a \bmod N, b \bmod N \in \mathbb{Z}/N\mathbb{Z}$ one defines

$$(a \bmod N) + (b \bmod N) = (r_1 + r_2) \bmod N$$

and

$$(a \bmod N) \cdot (b \bmod N) = r_1 r_2 \bmod N,$$

with r_1 an arbitrary element of $a \bmod N$ and similarly r_2 arbitrary in $b \bmod N$. Theorem II.1.5 shows that the resulting residue classes are independent of the choice of r_1, r_2.

II.1.7 Remark. If one chooses $r_1 = a$ and $r_2 = b$ in the above definition (which is allowed, since a, b are elements of $a \bmod N, b \bmod N$, respectively), then the definition reads $\overline{a} + \overline{b} = \overline{a + b}$ and $\overline{a} \cdot \overline{b} = \overline{ab}$.

II.1.8 Example. Take $N = 17$. It holds that $\overline{-1} = \overline{67}$, since the difference of -1 and 67 is divisible by 17. Hence also $\overline{1} = \overline{(-1)(-1)} = \overline{-1} \cdot \overline{-1} = \overline{67} \cdot \overline{67} = \overline{67^2}$. Apparently, 67^2 and 1 yield the same remainder upon division by 17, or in other words, $67^2 - 1$ is divisible by 17. (Of course this could be seen without the aid of residue classes: $67^2 - 1 = (67 + 1)(67 - 1)$.) ▬■

Since we can add and multiply residue classes modulo N, we can also raise them to a positive power n.

II.1.9 Definition. For a natural number n the nth power of a residue class $\overline{a}$, notation $\overline{a}^n$, is inductively defined as follows. $\overline{a}^1 = \overline{a}$, and if for $n \geq 1$ we have defined $\overline{a}^n$, then $\overline{a}^{n+1} = \overline{a}^n \cdot \overline{a}$.

It holds that $\overline{a}^m = \overline{a^m}$ and $\overline{a}^{n+m} = \overline{a}^n \cdot \overline{a}^m$, as is easily verified using mathematical induction w.r.t. m. Moreover $\overline{ab}^m = \overline{(ab)^m} = \overline{a^m b^m} = \overline{a^m} \cdot \overline{b^m} = \overline{a}^m \cdot \overline{b}^m$.

II.1.10 Example. To illustrate the use of these definitions, we will show that $2^{1000}+1$ is divisible by 257. Write $\overline{a}$ for the residue class of a modulo 257. Then

$$\overline{2^{1000}} = \overline{(2^8)^{125}} = \overline{256}^{125} = \overline{-1}^{125} = \overline{-1}.$$

Since 2^{1000} and -1 yield the same residue class modulo 257, their difference is divisible by 257 which is what we wanted to show. Note that $2^{1000}+1$ has 302 decimal digits, so to check the asserted divisibility using a simple division by 257 is quite elaborate. ∎

II.1.11 Example. We calculate the last two decimal digits of 2^{1000}. This is the same as finding the remainder of 2^{1000} upon division by 100. In $\mathbb{Z}/100\mathbb{Z}$ we have

$$\overline{16}^6 = \overline{4}^6 \cdot \overline{4}^6 = \overline{4096} \cdot \overline{4096} = \overline{-4} \cdot \overline{-4} = \overline{16},$$

since $4^6 = 2^{12} = 4096$. Furthermore $1000 = 4 \cdot 250$ and $250 = 6 \cdot 41 + 4$ and $41 = 6 \cdot 6 + 5$, so

$$\begin{aligned} \overline{2^{1000}} = \overline{2}^{250} &= \overline{16}^4 \cdot (\overline{16}^6)^{41} \\ &= \overline{16}^4 \cdot \overline{16}^{41} \qquad = \overline{16}^4 \cdot (\overline{16}^6)^6 \cdot \overline{16}^5 \\ &\qquad\qquad\quad = \overline{16}^4 \cdot \overline{16} \cdot \overline{16}^5 \\ &\qquad\qquad\quad = \overline{16}^4 \cdot \overline{16}^6 = \overline{16}^4 \cdot \overline{16} \\ &\qquad\qquad\quad = \overline{16}^5 = (\overline{2^{10}})^2 = \overline{24}^2 = \overline{76}. \end{aligned}$$

Hence the requested 2 decimals are 76. ∎

II.2 Units modulo N

II.2.1 Definition. A residue class a mod N is called a *unit modulo* N if a residue class b mod N exists such that $(a \bmod N) \cdot (b \bmod N) = 1 \bmod N$.

The subset of $\mathbb{Z}/N\mathbb{Z}$ consisting of all units modulo N is denoted as $(\mathbb{Z}/N\mathbb{Z})^\times$.

II.2.2 Example. Take $N = 12$. Then $\mathbb{Z}/12\mathbb{Z} = \{\overline{0}, \overline{1}, \ldots, \overline{10}, \overline{11}\}$. For each of these classes we check whether it is in $(\mathbb{Z}/12\mathbb{Z})^\times$. If $a, b \in \mathbb{Z}$ and $\overline{a} \cdot \overline{b} = \overline{1}$, this means that $ab = 1 + 12k$ for certain $k \in \mathbb{Z}$. In particular, if $\overline{a}$ is a unit modulo 12, then a is not divisible by 2, nor by 3. Hence

$$(\mathbb{Z}/12\mathbb{Z})^\times \subset \{\overline{1}, \overline{5}, \overline{7}, \overline{11}\}.$$

Since $\overline{1}^2 = \overline{5}^2 = \overline{7}^2 = \overline{11}^2 = \overline{1}$, the four given classes are indeed units modulo 12. So

$$(\mathbb{Z}/12\mathbb{Z})^\times = \{\overline{1}, \overline{5}, \overline{7}, \overline{11}\}.$$

∎

We present a simple criterion for finding the units modulo N.

II.2.3 Theorem. *Let $a \in \mathbb{Z}$. Then $a \bmod N \in (\mathbb{Z}/N\mathbb{Z})^\times$ if and only if $\gcd(a,N) = 1$.*

Proof. Let $a \in \mathbb{Z}$. The assertion '$a \bmod N \in (\mathbb{Z}/N\mathbb{Z})^\times$' by definition means that $b \in \mathbb{Z}$ exists with $(ab) \bmod N = (a \bmod N)\cdot(b \bmod N) = 1 \bmod N$. This is equivalent with the existence of $q, b \in \mathbb{Z}$ such that $ab - 1 = Nq$. Written differently: $ab - Nq = 1$. In Corollary I.1.14 it was shown that such integers exist precisely when $\gcd(a,N) = 1$. ∎

II.2.4 Definition. (The Euler Phi function; Leonhard Euler, Swiss mathematician, 1707–1783) The number of elements of $(\mathbb{Z}/N\mathbb{Z})^\times$ is denoted $\varphi(N)$.

II.2.5 Corollary. *$\varphi(N)$ equals the number of integers $a \in \mathbb{Z}$ with $1 \le a \le N$ and $\gcd(a,N) = 1$. In particular a positive integer p is prime if and only if $\varphi(p) = p - 1$.*

Proof. This is immediate from the definitions and from Theorem II.2.3. ∎

We list some properties of $(\mathbb{Z}/N\mathbb{Z})^\times$.

II.2.6 Theorem. *1. If $a \bmod N$ and $b \bmod N$ are units modulo N, then so is their product.*
2. If $a \bmod N \in (\mathbb{Z}/N\mathbb{Z})^\times$, then a residue class $b \bmod N$ such that $(a \bmod N)\cdot(b \bmod N) = 1 \bmod N$ is also a unit modulo N.
3. For each $a \bmod N \in (\mathbb{Z}/N\mathbb{Z})^\times$ there is a unique class $b \bmod N \in (\mathbb{Z}/N\mathbb{Z})^\times$ with $(a \bmod N)\cdot(b \bmod N) = 1 \bmod N$.

Proof. 1: By Theorem II.2.3 the assertion is equivalent with: if both $\gcd(a,N) = 1$ and $\gcd(b,N) = 1$, then also $\gcd(ab,N) = 1$. To see why the latter implication holds, assume $\gcd(ab,N) \neq 1$. Then a prime p dividing $\gcd(ab,N)$ exists. This prime divides N and ab, hence by Theorem I.2.3 we have $p|a$ or $p|b$. This contradicts $\gcd(a,N) = \gcd(b,N) = 1$.

Alternative proof: $\overline{a}, \overline{b}$ are units modulo N, so $\overline{c}, \overline{d}$ exist with $\overline{c}\cdot\overline{a} = \overline{db} = \overline{1}$. Put $\overline{e} = \overline{d}\cdot\overline{c}$, then $\overline{e}\cdot\overline{ab} = \overline{dcab} = \overline{d}\cdot\overline{ca}\cdot\overline{b} = \overline{db} = \overline{1}$, hence $\overline{ab} \in (\mathbb{Z}/N\mathbb{Z})^\times$.
2: This is immediate from the fact that $\overline{a}\cdot\overline{b} = \overline{ab} = \overline{ba} = \overline{b}\cdot\overline{a}$.
3: If $\overline{a}\cdot\overline{b_1} = \overline{1} = \overline{a}\cdot\overline{b_2}$, then also $\overline{b_1} = \overline{b_1}\cdot\overline{1} = \overline{b_1}\cdot\overline{ab_2} = \overline{b_1ab_2} = \overline{ab_1}\cdot\overline{b_2} = \overline{b_2}$. This proves Theorem II.2.6. ∎

II.2.7 Definition. For $\overline{a} \in (\mathbb{Z}/N\mathbb{Z})^\times$ the unique $\overline{b} \in (\mathbb{Z}/N\mathbb{Z})^\times$ with the property $\overline{a}\cdot\overline{b} = \overline{1}$ is called the *inverse* of $\overline{a}$, notation $\overline{a}^{-1}$.

II.2.8 Remark. For $a \in \mathbb{Z}$ with $a \bmod N$ a unit modulo N, the inverse of $a \bmod N$ may be found using the Euclidean algorithm. Namely $a \bmod N$ being a unit implies $\gcd(a,N) = 1$. So $x, y \in \mathbb{Z}$ exist with $xa + yN = 1$, and any such x satisfies $\overline{x}\cdot\overline{a} = \overline{1}$, in other words, $x \bmod N$ is the inverse of $a \bmod N$.

The most important operations we did so far on residue classes modulo N, are besides addition, subtraction and multiplication, the operations exponentiation (=raising to a positive power) and in the special case of units modulo N, taking the inverse.

In computer algebra systems such as MAGMA, Maple, Mathematica, and PARI, and even in WolframAlpha, standard routines are implemented for the mentioned operations. For example in Maple this may look as follows:

```
100^(-1) mod 420001;
7 &^ (420!) mod 100;
```

The symbol & in the second line makes sure, that Maple does not first raise 7 to the power 420!, and subsequently find the remainder of the result upon division by 100. Instead, a far more efficient way to obtain the answer is used.

II.2.9 Example. It holds that $(13 \bmod 37)^{-1} = 20 \bmod 37$ (check this!). ∎

II.2.10 Theorem. *For all* $a \bmod N \in (\mathbb{Z}/N\mathbb{Z})^\times$ *one has* $(a \bmod N)^{\varphi(N)} = 1 \bmod N$.

This result is due to Euler. An alternative way to formulate the theorem is: if $a, N \in \mathbb{Z}$ satisfy $N > 0$ and $\gcd(a,N) = 1$, then $N | a^{\varphi(N)} - 1$. The theorem and some of its consequences are nowadays used for example in cryptography and in testing whether a (very large) integer is prime. We return to this below.

Proof. Write $(\mathbb{Z}/N\mathbb{Z})^\times = \{\overline{a_1}, \ldots, \overline{a_{\varphi(N)}}\}$. In Theorem II.2.6 we saw that a product of units modulo N is a unit as well, hence

$$\epsilon := \overline{a_1} \cdot \overline{a_2} \cdot \ldots \cdot \overline{a_{\varphi(N)}}$$

is a unit modulo N. Consider the map 'multiplication by $\overline{a}$'. This map sends $(\mathbb{Z}/N\mathbb{Z})^\times$ to itself, since $\overline{a}$ is a unit. We will now show that this map is a bijection from $(\mathbb{Z}/N\mathbb{Z})^\times$ to itself. Indeed, $\overline{a} \cdot \overline{b} = \overline{a} \cdot \overline{c}$ implies, by multiplying both products with the inverse of $\overline{a}$, that $\overline{b} = \overline{c}$. Hence the map is injective. This means that distinct elements are mapped to distinct elements, and therefore the image consists of equally many elements as the source, namely $\varphi(N)$. As a consequence the map is surjective as well. This means

$$(\mathbb{Z}/N\mathbb{Z})^\times = \{\overline{a} \cdot \overline{a_1}, \overline{a} \cdot \overline{a_2}, \ldots, \overline{a} \cdot \overline{a_{\varphi(N)}}\}.$$

As we saw, multiplying these elements yields ϵ. On the other hand this product equals

$$(\overline{a} \cdot \overline{a_1}) \cdot (\overline{a} \cdot \overline{a_2}) \cdot \ldots \cdot (\overline{a} \cdot \overline{a_{\varphi(N)}}) = \overline{a}^{\varphi(N)} \cdot \overline{a_1} \cdot \overline{a_2} \cdot \ldots \cdot \overline{a_{\varphi(N)}} = \overline{a}^{\varphi(N)} \cdot \epsilon.$$

This shows $\epsilon = \overline{a}^{\varphi(N)} \cdot \epsilon$, and multiplying by the inverse of ϵ now gives $\overline{a}^{\varphi(N)} = \overline{1}$, which is what we wanted to prove. ∎

II.2.11 Corollary. (Fermat's little theorem; Pierre de Fermat, French amateur mathematician, 1601–1665) *Is* p *prime, then* $(a \bmod p)^p = a \bmod p$ *for every* $a \in \mathbb{Z}$.

Proof. If $a \bmod p$ is not a unit modulo p, then p being prime implies $a \bmod p = 0 \bmod p$, and in this case the corollary is clear. Now note that since p is prime, $\varphi(p) = p - 1$ as follows from Corollary II.2.5. Hence for $a \bmod p$ a unit one finds using Theorem II.2.10 that $(a \bmod p)^p = (a \bmod p) \cdot (a \bmod p)^{p-1} = (a \bmod p) \cdot \overline{1} = a \bmod p$. ∎

Fermat's little theorem yields and efficient and simple criterion which can be used to check that certain large integers are not prime. Namely, to test whether N is prime, take (for example) $a = 2$, and calculate $(2 \bmod N)^N$. If the result is not $2 \bmod N$, then Fermat's little theorem allows one to conclude that N is not prime. Here the exponentiation can be done by computing in the order of $\log(N)$ multiplications/divisions with remainder. Moreover, this involves only integers between 0 and N. Hence this 'compositeness test' is much faster than simply testing whether N has some divisor between 1 and $\sqrt{N}$! However as a disadvantage, our algorithm yields not as much information: for example, if one concludes from the algorithm that N is not prime, this does not provide any relevant information concerning possible divisors of N. A much worse observation is: there exist composite integers such as $341 = 11 \cdot 31$, which has the property $(2 \bmod 341)^{341} = 2 \bmod 341$. In this example one can consider $3 \bmod 341$ instead of $2 \bmod 341$, and thereby test that

341 is not prime. In this respect, a real nuisance are the so-called Carmichael numbers. These are composite positive integers N with the property that every $a \in \mathbb{Z}$ satisfies $(a \bmod N)^N = a \bmod N$. The smallest Carmichael number is $N = 561 = 3 \cdot 11 \cdot 17$. In 1992 Alford, Granville and Pomerance proved that infinitely many such Carmichael numbers exist.

II.3 The Chinese remainder theorem

To check that for example $N = 561 = 3 \cdot 11 \cdot 17$ indeed has the aforementioned property, namely that every integer a satisfies $561 | a^{561} - a$, it is natural to test, instead of divisibility by $561 = 3 \cdot 11 \cdot 17$, divisibility by 3, 11, and 17. In this way the property can be checked rather simply, as follows: in case is a coprime to 3, Theorem II.2.10 shows that $(a \bmod 3)^2 = 1 \bmod 3$, hence any *even* exponent $m = 2k$ satisfies $(a \bmod 3)^m = (1 \bmod 3)^k = 1 \bmod 3$, i.e., $3 | a^{2k} - 1$. Multiplying by a then shows $3 | a^{2k+1} - a$. This evidently holds as well for any a which is divisible by 3, so for all $a \in \mathbb{Z}$.

In the same way one shows that any exponent $m = 10\ell$ which is a multiple of 10 satisfies $11 | a^{10\ell} - 1$ (provided a is coprime to 11). As a consequence, every $a \in \mathbb{Z}$ has the property $11 | a^{10\ell+1} - a$. Finally, a similar reasoning shows $17 | a^{16n+1} - a$ for all $a \in \mathbb{Z}$. Combining the three divisibility properties above, one concludes that any exponent m equal to 1 plus a multiple of each of 16, 10, and 2, in other words m equals 1 plus a multiple of $\operatorname{lcm}(16, 10, 2) = 80$, satisfies $3 \cdot 11 \cdot 17 = 561 | a^m - a$ for all $a \in \mathbb{Z}$. In particular the exponent $m = 561 = 1 + 80 \cdot 7$ has the required form, so $561 | a^{561} - a$.

The above result was obtained by combining two ingredients: arithmetic modulo a prime p (implying the necessary special case of Theorem II.2.10), and deducing from divisibility by primes divisibility by their product. This last ingredient we now consider more generally; in particular, not only for primes.

II.3.1 Lemma. *Suppose $N, M \in \mathbb{Z}$ are positive. The rule 'send the residue class of a modulo N to the residue class of a modulo M' yields a well defined map: $\mathbb{Z}/N\mathbb{Z} \longrightarrow \mathbb{Z}/M\mathbb{Z}$ if and only if $M | N$.*

Proof. We are asked to examine under which condition(s) the description above defines a map. Consider any residue class in $\mathbb{Z}/N\mathbb{Z}$; say, this is the residue class of $a \in \mathbb{Z}$. It is also the residue class of $a + N, a + 2N, a - N$, and more generally of $b \in \mathbb{Z}$ for any b with $N | a - b$. The desired map would send this class to $a \bmod M$, but also, for any $b \in \mathbb{Z}$ as above, to $b \bmod M$. So the assignment is well defined, precisely when all $a, b \in \mathbb{Z}$ such that $N | a - b$ satisfy $a \bmod M = b \bmod M$. In other words: $N | a - b$ should imply $M | a - b$. We need this condition for all $a, b \in \mathbb{Z}$. Taking $a = N$ and $b = 0$ shows the necessary condition $M | N$. Vice versa, is $M | N$, then for all $a, b \in \mathbb{Z}$ with $N | a - b$ one obtains $M | N | a - b$, hence in particular $M | a - b$. This proves the lemma. ∎

II.3.2 Remark. The lemma above may appear somewhat strange. However, it reveals a very essential property of modular arithmetic. Namely, residue classes are sets, and if one picks an element from such a set and performs certain operations on it, the final result may very well change if a different element from the same residue class was chosen.

II.3.3 Example. $a \bmod 4 \mapsto a \bmod 2$ defines a map from $\mathbb{Z}/4\mathbb{Z}$ to $\mathbb{Z}/2\mathbb{Z}$. The image of each of 1 mod 4 and 3 mod 4 is 1 mod 2, and both 0 mod 4 and 2 mod 4 have image 0 mod 2.

On the other hand $a \bmod 2 \mapsto a \bmod 4$ is *not* well defined. For example, 1 mod 2 and 3 mod 2 are the same residue class, but 1 mod 4 and 3 mod 4 differ.

Stated differently, this example shows: if we know the remainder upon division by 4 of some integer, then we also know its remainder upon division by 2. On the other hand, given the remainder upon division by 2 one cannot conclude what is the remainder upon division by 4. ▬

In order to formulate the main results of this section, we briefly recall a notation from basic set theory. Given sets V en W, one denotes the cartesian product of V and W as $V \times W$. By definition this consists of all ordered pairs consisting of an element from V followed by an element from W:

$$V \times W = \{(v,w) \mid v \in V \text{ and } w \in W\}.$$

II.3.4 Theorem. (The Chinese remainder theorem)
Let N, M be positive integers with $\gcd(N,M) = 1$. The assignment

$$a \bmod NM \mapsto (a \bmod N, a \bmod M) : \mathbb{Z}/NM\mathbb{Z} \longrightarrow \mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/M\mathbb{Z}$$

defines a well-defined map. This map is bijective.

Moreover it maps $(\mathbb{Z}/NM\mathbb{Z})^\times$ to $(\mathbb{Z}/N\mathbb{Z})^\times \times (\mathbb{Z}/M\mathbb{Z})^\times$, and this is a bijection as well.

Proof. The assignment is well defined by Lemma II.3.1. Next, we show that the given map is injective. If $a, b \in \mathbb{Z}$ and $(a \bmod N, a \bmod M) = (b \bmod N, b \bmod M)$, this means by definition $N|a-b$ and $M|a-b$. Now Corollary I.2.9 implies that $\operatorname{lcm}(N,M)|a-b$. Furthermore, since $\gcd(N,M) = 1$, from the same corollary one obtains $NM = \gcd(N,M) \cdot \operatorname{lcm}(N,M) = \operatorname{lcm}(N,M)$. Conclusion: $NM|a-b$, i.e., $a \bmod NM = b \bmod NM$. The map is therefore injective, and since both $\mathbb{Z}/NM\mathbb{Z}$ and $\mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/M\mathbb{Z}$ consist of NM elements, the map is then surjective as well.

If $a \in \mathbb{Z}$ satisfies $a \bmod NM \in (\mathbb{Z}/NM\mathbb{Z})^\times$, then $\gcd(a,NM) = 1$, hence in particular also $\gcd(a,N) = 1 = \gcd(a,M)$. This means precisely that $(a \bmod N, a \bmod M) \in (\mathbb{Z}/N\mathbb{Z})^\times \times (\mathbb{Z}/M\mathbb{Z})^\times$. So our map sends $(\mathbb{Z}/NM\mathbb{Z})^\times$ to $(\mathbb{Z}/N\mathbb{Z})^\times \times (\mathbb{Z}/M\mathbb{Z})^\times$.

Vice versa, if $(a \bmod N, b \bmod M) \in (\mathbb{Z}/N\mathbb{Z})^\times \times (\mathbb{Z}/M\mathbb{Z})^\times$ for certain $a, b \in \mathbb{Z}$, then $\gcd(a,N) = 1 = \gcd(b,N)$. Since our map is surjective, $c \in \mathbb{Z}$ exists with $(c \bmod N, c \bmod M) = (a \bmod N, b \bmod M)$. This means $c = q_1 N + a$ and $c = q_2 M + b$, so by Lemma I.1.8 one concludes $\gcd(c,N) = \gcd(a,N) = 1$ and $\gcd(c,M) = \gcd(b,M) = 1$. As a consequence $\gcd(c,NM) = 1$, i.e., $c \bmod NM \in (\mathbb{Z}/NM\mathbb{Z})^\times$. So restricting our map to $(\mathbb{Z}/NM\mathbb{Z})^\times$ one obtains as image $(\mathbb{Z}/N\mathbb{Z})^\times \times (\mathbb{Z}/M\mathbb{Z})^\times$, and evidently this restriction is injective as well. This proves the Chinese remainder theorem. ■

II.3.5 Example. The Chinese remainder theorem can be viewed as a solvability criterion for systems of simultaneous congruences: given a, b, N, M one asks for $x \in \mathbb{Z}$ satisfying both $x \equiv a \bmod N$ and $x \equiv b \bmod M$. The theorem states that in case $\gcd(N,M) = 1$, then a solution x exists, and moreover the set of all solutions is a residue class modulo NM.

Example: we find all $x \in \mathbb{Z}$ with $x \equiv 4 \bmod 9$ and $x \equiv 5 \bmod 11$. Such x necessarily have the form $x = 4 + 9y$, with $y \in \mathbb{Z}$. Moreover we demand $x = 4 + 9y \equiv 5 \bmod 11$, i.e., $9y \equiv 1 \bmod 11$. This means precisely that $y \bmod 11$ is the inverse of 9 mod 11 in $(\mathbb{Z}/11\mathbb{Z})^\times$, so $y \bmod 11 = 5 \bmod 11$. So $y = 5 + 11z$ hence $x = 4 + 9(5 + 11z) = 49 + 99z$ with $z \in \mathbb{Z}$ arbitrary. Stated differently: the solution equals the residue class of 49 modulo 99. The only integer between 1 and 99 with remainder 4 upon division by 9 and remainder 5 upon division by 11, is therefore 49. ▬

II.3.6 Remark. Using mathematical induction w.r.t. n one may generalize the Chinese remainder theorem as follows. Suppose $N_1, \ldots, N_n$ are positive integers and

$\gcd(N_i, N_j) = 1$ for all pairs i, j with $1 \le i < j \le n$. Then

$$\mathbb{Z}/N_1 \ldots N_n\mathbb{Z} \longrightarrow \mathbb{Z}/N_1\mathbb{Z} \times \mathbb{Z}/N_2\mathbb{Z} \times \ldots \times \mathbb{Z}/N_n\mathbb{Z}$$

given by $a \bmod N_1 \ldots N_n \mapsto (a \bmod N_1, \ldots, a \bmod N_n)$ defines a bijection. The same holds if one restricts the map to units.

Example: $7, 11$, and 13 are pairwise coprime, with product $7 \cdot 11 \cdot 13 = 1001$. For every triple of integers $a, b, c \in \mathbb{Z}$ a unique $x \in \mathbb{Z}$ with $0 \le x \le 1000$ and $x \equiv a \bmod 7$ and $x \equiv b \bmod 11$ and $x \equiv c \bmod 13$ exists. Try for yourself to find this x for certain triples a, b, c.

II.3.7 Corollary. *Euler's φ-function has the property $\varphi(NM) = \varphi(N) \cdot \varphi(M)$ for all positive coprime integers N, M.*

Proof. By definition $\varphi(n)$ equals the number of elements of $(\mathbb{Z}/n\mathbb{Z})^\times$. Hence the assertion is a direct consequence of Theorem II.3.4, using the fact that for finite sets V, W the number of elements of $V \times W$ equals the product of the numbers of elements of V and of W. ∎

Using this corollary we will derive a formula for $\varphi(n)$ in terms of the prime factorization of n. This uses the next result.

II.3.8 Lemma. *For p prime and k an integer ≥ 1 we have*

$$\varphi(p^k) = (p-1)p^{k-1} = p^k - p^{k-1}.$$

Proof. We know that $\varphi(p^k)$ equals the number of integers a with $0 \le a \le p^k - 1$ and $\gcd(a, p^k) = 1$. The given interval contains precisely p^k integers, and any integer is *not* coprime to p^k, precisely when it is divisible by p. The integers in our interval which are divisible by p are $0 \cdot p, 1 \cdot p, \ldots, m \cdot p$, with m the largest integer smaller than p^{k-1}. The interval therefore contains p^{k-1} integers divisible by p, hence $\varphi(p^k) = p^k - p^{k-1}$. ∎

II.3.9 Theorem. *The Euler φ-function can be computed, for $n \ge 2$, using the formula*

$$\varphi(n) = \prod_{p|n} (p-1)p^{v_p(n)-1} = n \prod_{p|n} (1 - \frac{1}{p})$$

where the product is taken over the prime divisors of n.

Proof. The second equality follows using $n = \prod_{p|n} p^{v_p(n)}$.

For the first equality we will use mathematical induction w.r.t. N, showing that the formula holds for every n with $2 \le n \le N$. For $N = 2$ this is clear. Assuming it for $N \ge 2$, then to verify the case $N+1$ we only need to show the formula for $n = N+1$. Is this n a power of a prime, then we are done by Lemma II.3.8. If not, write $n = p^{v_p(n)} \cdot n'$, with $2 \le n' \le N$. Then $\gcd(p^{v_p(n)}, n') = 1$, so by Corollary II.3.7 we have $\varphi(n) = \varphi(p^{v_p(n)})\varphi(n')$. Applying the induction hypothesis to n' then implies the formula for n (using that $v_q(n) = v_q(n')$ for every prime $q \ne p$). ∎

II.3.10 Example. $\varphi(1000000) = 2^5 \cdot 4 \cdot 5^5 = 400000$. So 400000 positive odd integers below one million exist with last decimal digit different from 5. ∎

II.4 Exercises

1. Prove that for every odd $n \in \mathbb{Z}$ the congruence $n^2 \equiv 1 \bmod 8$ holds, and for every odd prime $p \neq 3$ we even have $p^2 \equiv 1 \bmod 24$.
2. Given an integer $n = \sum a_i 10^i$ (with all $a_i \in \mathbb{Z}$).
   (a) Show: for $p = 2$ and for $p = 5$ it holds that $p|n$ if and only if $p|a_0$.
   (b) Show: for $m = 3$ and for $m = 9$ one has $m|n$ if and only if $m|\sum a_i$.
   (c) Prove that $11|n$ if and only if $11|\sum(-1)^i a_i$.
3. Determine the inverse of $\overline{100}$ in $(\mathbb{Z}/257\mathbb{Z})^\times$.
4. Show that $2^{341} \equiv 2 \bmod 341$. Is 341 prime? Find an integer between 0 and 341 that is congruent to $3^{341} \bmod 341$.
5. Show that every $n \in \mathbb{Z}$ satisfies $n^{13} \equiv n \bmod 2730$.
6. Find the remainder upon dividing $(177 + 10^{15})^{116}$ by $1003 = 17 \times 59$.
7. Find all integers which leave a remainder 3 upon division by 7, remainder 6 upon division by 11, and remainder 1 upon division by 13.
8. (a) Determine for $n = 4$ the residue class $(n-1)! \bmod 4$.
   (b) Show that if $n > 4$ is not prime, then $(n-1)! \equiv 0 \bmod n$.
   (c) Now suppose $n = p$ is prime. Find all residue classes $a \bmod p \in (\mathbb{Z}/p\mathbb{Z})^\times$ satisfying $(a \bmod p)^{-1} = a \bmod p$.
   (d) Show for $n = p$ prime that $(n-1)! \equiv -1 \bmod n$.
   (e) Show that $n \geq 2$ is prime if and only if $(n-1)! \equiv -1 \bmod n$. Is this a practical way to test primality?
9. (a) Prove that a Carmichael number is not divisible by any square > 1.
   (b) Prove that if $n = p_1 \cdot \ldots \cdot p_t$ is a product of $t > 1$ distinct primes, and $p_i - 1|n-1$ for every i, then n is a Carmichael number.
10. Find all positive integers n satisfying $\varphi(n) = 24$. Answer the analogous question for $\varphi(n) = 14$.
11. Let $N, a \in \mathbb{Z}$ with $N > 0$.
    (a) Prove that $\overline{a} \in (\mathbb{Z}/N\mathbb{Z})^\times$ if and only if $\overline{-a} \in (\mathbb{Z}/N\mathbb{Z})^\times$.
    (b) Which residue classes $a \bmod N$ satisfy $a \bmod N = -a \bmod N$? (Distinguish the cases N even and N odd.)
    (c) Show that $\varphi(N)$ is even for all $N \geq 3$.
12. Show that if $p_1, \ldots, p_t$ are the smallest t primes, and $n_j = p_1 \cdots p_t - p_1 \cdots p_t / p_j$, then $\varphi(n_j) = \varphi(n_k)$ for $1 \leq j, k \leq t$. Conclude from this that for fixed m, the equation $\varphi(x) = m$ may have arbitrary many solutions.
13. Find all n such that $\varphi(n)|n$.

III GROUPS AND HOMOMORPHISMS

A number of properties $\mathbb{Z}$ has with respect to addition, one finds in $(\mathbb{Z}/N\mathbb{Z})^\times$ for the multiplication. Below we will see many more sets equipped with an operation, all sharing the same properties. It is typical for (abstract) algebra to capture such a phenomenon in a definition. Instead of dealing with all separate cases one by one, this makes it possible to prove things at once for all examples satisfying the definition. We saw a similar situation in linear algebra: after abstractly introducing the notion 'vector space over a field', one deduces a range of properties not only for well known spaces such as $\mathbb{R}^2$ and $\mathbb{R}^3$, but also for planes and hyperplanes containing the origin in $\mathbb{R}^n$, for function spaces, spaces of polynomials over the complex numbers, spaces of sequences, matrices, et cetera.

For the topic 'groups' which will be discussed here, the first to formulate an abstract definition was the German mathematician W.F.A. von Dyck (1856–1934). Algebra courses starting from abstract definitions of this kind were presented in Göttingen around 1920, notably by the famous female mathematician Emmy Noether (1882–1935). In her audience was a young student from Amsterdam, B.L. van der Waerden. He extended his algebra knowledge with help of Emil Artin (1898–1962) in Hamburg. In 1928, only 25 years old, Van der Waerden was appointed mathematics professor in Groningen where he wrote what is probably the most influential textbook on abstract algebra to date. It appeared in 1930 and completely adopts the abstract definition/theorem/proof style. The book made Van der Waerden, who died in 1996, world famous. Due to Noether's and Artin's lectures and Van der Waerden's recording of this, abstract algebra is still taught all over the world essentially exclusively in this style.

III.1 Groups

III.1.1 Definition. A *group* is a triple $(G,\cdot,e)$ with G a set, $e\in G$, and $\cdot$ a map from $G\times G$ to G, which we write as $(x,y)\mapsto x\cdot y$, satisfying

G1 (associativity) For all $x,y,z\in G$ we have $(x\cdot y)\cdot z = x\cdot(y\cdot z)$.
G2 (unit element) For all $x\in G$ we have $e\cdot x = x = x\cdot e$.
G3 (inverses) For all $x\in G$ a $y\in G$ exists such that $x\cdot y = e = y\cdot x$.

A group $(G,\cdot,e)$ is called *commutative* or (in honour of the Norwegian mathematician Niels Henrik Abel, 1802–1829) *abelian*, if moreover

G4 For all $x,y\in G$ we have $x\cdot y = y\cdot x$.

III.1.2 Remark. Instead of $(G,\cdot,e)$ one usually simply writes G, assuming that the map $G\times G\to G$ and the element $e\in G$ are clear from the context. The map $\cdot$ is called multiplication on G or group law on G. Instead of $x\cdot y$, depending on the context other notations are used, such as $x\circ y$ or $x*y$ or $x\times y$ or $x+y$ or even xy.

III.1.3 Example. $(\mathbb{Z},+,0)$ is a group, as well as $(\mathbb{Z}/N\mathbb{Z},+,\overline{0})$ and $((\mathbb{Z}/N\mathbb{Z})^\times,\cdot,\overline{1})$. These are examples of commutative groups. We already know many more commutative groups: if $G=V$ a vector space (over some field) with vector addition '+' and zero vector $0\in V$, then $(V,+,0)$ is an abelian group. ∎

III.1.4 Example. The set of invertible $n\times n$ matrices with coefficients in $F=\mathbb{R}$ or $F=\mathbb{C}$ or $F=\mathbb{Q}$ (or more generally: in a *field* F) becomes a group, when one takes as group law the multiplication of matrices, and as unit the unit matrix. This group is denoted as $\mathrm{GL}_n(F)$. For $n\geq 2$ this group is *not* commutative, because (for example)

$$\begin{pmatrix} 1 & 0 & \cdots & \cdots & 0\\ 1 & 1 & 0 & \cdots & 0\\ 0 & 0 & \cdots & \cdots & \vdots\\ \vdots & & & & \vdots\\ 0 & & \cdots & & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 & \cdots & 0\\ 0 & 1 & 0 & \cdots & 0\\ 0 & 0 & \cdots & \cdots & \vdots\\ \vdots & & & & \vdots\\ 0 & & \cdots & & 1 \end{pmatrix}$$
$$\neq$$
$$\begin{pmatrix} 1 & 1 & 0 & \cdots & 0\\ 0 & 1 & 0 & \cdots & 0\\ 0 & 0 & \cdots & \cdots & \vdots\\ \vdots & & & & \vdots\\ 0 & & \cdots & & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & \cdots & \cdots & 0\\ 1 & 1 & 0 & \cdots & 0\\ 0 & 0 & \cdots & \cdots & \vdots\\ \vdots & & & & \vdots\\ 0 & & \cdots & & 1 \end{pmatrix}.$$

∎

Given elements $a_1,a_2,\ldots,a_n$ in a group G, their product $a_1a_2\ldots a_n$ is inductively w.r.t. n defined as follows. For $n=2$ it is just the product in the group. If the product is defined for $n-1\geq 2$, then $a_1a_2\ldots a_n=(a_1\ldots a_{n-1})\cdot a_n$. Using associativity and induction w.r.t. n it turns out that $(a_1a_2\ldots a_k)\cdot(a_{k+1}\ldots a_n)=(a_1a_2\ldots a_n)$. This product of n factors is usually abbreviated as $\prod_{i=1}^n a_i$, and in case of an abelian group as $\sum_{i=1}^n a_i$. If all a_i are equal, say $a_i=a$, then we write $a^n=\prod_{i=1}^n a_i$. The property mentioned above, writing $\ell=n-k$, translates into $a^k\cdot a^\ell=a^{k+\ell}$. Note that in general exponentiation in a group behaves not quite as nice as exponentiation with e.g. integers. For example in the case of invertible matrices, in general the property $(AB)^2=A^2B^2$ does *not* hold (find an explicit example yourself).

We now present some elementary properties of groups.

III.1.5 Theorem. *Given is a group $(G,\cdot,e)$.*

1. *The only element $e'\in G$ such that $e'x=x$ for some $x\in G$, is $e'=e$. The same is true with the property $xe'=x$ for an $x\in G$.*
2. *For each $x\in G$ precisely one $y\in G$ with $xy=e=yx$ exists.*
3. *For any fixed $a\in G$, the map $x\mapsto ax$ is a bijection from G to itself. Similarly $x\mapsto xa$ is a bijection.*

Proof. 1: If $x\in G$ satisfies $e'x=x$, then multiplying on the right by a $y\in G$ with $xy=e$ (such y exists because of group property G3) shows $e'=e'e=e$; here the first equality follows from group property G2. The case $xe'=x$ is dealt with in a similar way, by multiplying on the left with $y\in G$ such that $yx=e$.
2: If $xy=e=yx$ and $xz=e=zx$, then $z=ze=z(xy)=(zx)y=ey=y$, so $y=z$.

3: Let $a \in G$. The map $x \mapsto ax$ is injective, for suppose $ax = ay$. Take $b \in G$ with $ba = e$, then $x = ex = (ba)x = b(ax) = b(ay) = (ba)y = ey = y$, so $x = y$. The map is surjective as well, because if $z \in G$, then take b as above and $x = bz$, then x is mapped to $ax = a(bz) = (ab)z = ez = z$, so z is in the image. The map 'multiply on the left by a ' is therefore both injective and surjective, hence bijective. The case 'multiply on the right by a' is completely analogous. ∎

III.1.6 Definition. Let $(G,\cdot,e)$ be a group and $x \in G$. The (by Theorem III.1.5 unique) element $y \in G$ such that $xy = e = yx$ is called the *inverse* of x in G. It is denoted by x^{-1}.

In case of an abelian group G with group law denoted as $+$, this inverse element is called the *opposite* of x in G, and it is denoted as $-x$.

III.1.7 Remark. To check that some element y in a group is the inverse of an element x, it suffices to verify that $xy = e$: namely, Theorem III.1.5(3.) implies that only one element in the group has this property, and by definition x^{-1} has the property. Similarly, it is enough to check that $yx = e$.

III.1.8 Corollary. Let G be a group and $a, a_1, a_2, \ldots, a_n \in G$.

1. $(a^{-1})^{-1} = a$.
2. $(a_1 \ldots a_n)^{-1} = a_n^{-1} \cdot a_{n-1}^{-1} \cdot \ldots \cdot a_1^{-1}$. (When taking an inverse, the order reverses!)
3. $(a^n)^{-1} = (a^{-1})^n$.

Proof. 1: Note that $aa^{-1} = e = a^{-1}a$, which says that a is the inverse of a^{-1}, i.e., $(a^{-1})^{-1} = a$.
2: We show this by induction w.r.t. n. For $n = 1$ the statement holds. If $n \geq 2$ and we assume $(a_1 \ldots a_{n-1})^{-1} = a_{n-1}^{-1} \cdot \ldots \cdot a_1^{-1}$, then

$$(a_n^{-1} \cdot a_{n-1}^{-1} \cdot \ldots \cdot a_1^{-1}) \cdot (a_1 \ldots a_n) = a_n^{-1}((a_{n-1}^{-1} \cdot \ldots \cdot a_1^{-1}) \cdot (a_1 \ldots a_{n-1}))a_n = a_n^{-1} e a_n = e$$

and similarly $(a_1 \ldots a_n) \cdot (a_n^{-1} \cdot a_{n-1}^{-1} \cdot \ldots \cdot a_1^{-1}) = e$. This finishes the proof.
3: This is the above statement with all a_i equal to a. ∎

If x is an element of a group G and $n \in \mathbb{Z}$, then we already defined x^n for positive n. For $n = 0$, by definition $x^0 = e$, and for negative n we define $x^n = (x^{-1})^{-n}$. Using what is proven above, we see that $x^n \cdot x^{-n} = e$, so x^{-n} is the inverse of x^n. Moreover, using mathematical induction w.r.t. $|n|$ it is not hard to verify that for $n, m \in \mathbb{Z}$ it holds that $x^{n+m} = x^n \cdot x^m$.

III.1.9 Definition. The multiplication table
One can describe a group G consisting of only finitely many elements by means of a table containing all results of multiplying pairs of elements of G. We represent this in a matrix $(a_{i,j})$. Position $a_{1,1}$ remains empty, or we could write the name of the group here. In the remainder of the first row we write the elements of G, and the same in the first column. In position $a_{i,j}$ (with $i, j \geq 2$) we put the product $a_{i,1} \cdot a_{1,j}$. (The product with an element from column one on the left, and an element from row one on the right! this order obviously makes a difference in the case of non-abelian groups.)

III.1.10 Example. Here is the multiplication table of $\mathbb{Z}/3\mathbb{Z}$:

$\mathbb{Z}/3\mathbb{Z}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{0}$	$\overline{1}$

∎

The fact that left multiplication by a fixed element is bijective, precisely means that in every row, all elements of the group appear exactly once (after the first position). In the same way right multiplication by a fixed element is bijective, and this says that in every column, from the second position onward, all elements occur once.

Testing whether a group is commutative means, in terms of the table that one verifies $a_{i,j} = a_{j,i}$ for all i, j. In other words, we need to check whether the matrix is symmetric.

III.1.11 Example. Here is the multiplication table of a non-abelian group consisting of 6 elements:

G	e	a	b	c	d	f
e	e	a	b	c	d	f
a	a	e	f	d	c	b
b	b	d	e	f	a	c
c	c	f	d	e	b	a
d	d	b	c	a	f	e
f	f	c	a	b	e	d

III.2 Subgroups

Analogous to the notion "linear subspace" in a vector space over a field, a similar notion in Group Theory occurs.

III.2.1 Definition. A *subgroup* H of a group $(G, \cdot, e)$ is a subset of G which, with the same element e and the same group law $\cdot$, is itself a group.

III.2.2 Example. In $(\mathbb{Z}, +, 0)$ the set of all even integers is a subgroup.

The set consisting of all non-negative integers is *not* a subgroup of $\mathbb{Z}$. Namely, with respect to the group law on $\mathbb{Z}$ (addition) not every element x in the subset satisfies property G3.

In the group $\mathrm{GL}_n(\mathbb{Q})$ the matrices with determinant 1 form a subgroup. This follows from the formula $\det(AB) = \det(A)\det(B)$ for $n \times n$-matrices A, B. This subgroup is denoted by $\mathrm{SL}_n(\mathbb{Q})$. The same conclusion holds if we replace $\mathbb{Q}$ by $\mathbb{R}$ or by $\mathbb{C}$ or more generally, by an arbitrary field F. In this case, the subgroup of matrices with determinant 1 is denoted $\mathrm{SL}_n(F)$.

To test efficiently whether a given subset of a group is a subgroup, one can use the following criterion.

III.2.3 Theorem. *Let $(G, \cdot, e)$ be a group and $H \subset G$. Then H is a subgroup if and only if*

H1 *$e \in H$;*
H2 *For all $x, y \in H$ also $x \cdot y \in H$;*
H3 *For all $x \in H$ also $x^{-1} \in H$.*

Proof. If H is a subgroup of $(G, \cdot, e)$, then by definition $(H, \cdot, e)$ is a group. This implies the properties H1, H2, and H3.

Vice versa, suppose that a subset H has properties H1, H2, and H3. We have to show that $(H, \cdot, e)$ is a group. H1 says that indeed $e \in H$, and H2 says that the restriction to H of the group law on G is indeed a map $H \times H \to H$. The triple

$(H,\cdot,e)$ satisfies G1 and G2, since these properties hold for all of G hence also for the subset H. Finally, because of H3 the triple $(H,\cdot,e)$ satisfies G3. ∎

III.2.4 Example. If G is a group and $x \in G$, then the set of all powers of x (positive as well as negative powers, and also x^0 which by definition equals e) is a subgroup. Namely, this set satisfies H1, H2, and H3. We denote the subgroup obtained in this way as $\langle x \rangle$; some texts also use the notation $x^{\mathbb{Z}}$. ▬

III.2.5 Example. In $(\mathbb{Z}/24\mathbb{Z})^\times$ we find various subgroups $\langle x \bmod 24 \rangle = \langle \overline{x} \rangle$, namely $\langle \overline{1} \rangle$ consisting of only one element, and $\langle \overline{5} \rangle, \langle \overline{7} \rangle, \langle \overline{11} \rangle, \langle \overline{13} \rangle, \langle \overline{17} \rangle, \langle \overline{19} \rangle, \langle \overline{23} \rangle$ each having precisely two elements.

Incidentally, $(\mathbb{Z}/24\mathbb{Z})^\times$ contains even more subgroups; for example also $\{\overline{\pm 1}, \overline{\pm x}\}$ for $x = 5, 7, 11$, each consisting of 4 elements. ▬

III.2.6 Example. We will now describe all subgroups of $(\mathbb{Z},+,0)$. To start with, $\{0\} = 0\mathbb{Z}$ is a subgroup. If H is a subgroup and $H \neq 0\mathbb{Z}$, then H contains an element $x \neq 0$. Since H is a group w.r.t. the usual addition, also $-x \in H$. So we may conclude that H contains at least one positive integer. The smallest positive integer contained in H we denote by a.

Claim: $H = \langle a \rangle = a\mathbb{Z}$. Namely, the second equality is simply the definition of $\langle a \rangle$ as given in Example III.2.4. To show that $H \supset \langle a \rangle$, one needs to verify that for every $n \in \mathbb{Z}$ it holds that $an \in H$. This can be done using mathematical induction w.r.t. $|n|$, using the properties H2 and H3 of H (work out the details yourself!). Vice versa, one has to show $H \subset a\mathbb{Z}$. Take and arbitrary $b \in H$. Let $d = \gcd(a,b)$. Then $d \in H$, because $x, y \in \mathbb{Z}$ exist with $d = ax + by$. Now $ax \in H$ (this argument is explained above) and with $b \in H$ also $by \in H$. Property H2 therefore implies $d = ax + by \in H$. We have $1 \leq d \leq a$, so because a is by definition the smallest positive integer in H, it follows that $d = a$. This implies $a = d|b$, i.e., $b \in a\mathbb{Z}$. Having verified both inclusions we conclude $H = \langle a \rangle$.

An arbitrary subgroup of $\mathbb{Z}$ is therefore of the form $a\mathbb{Z}$, with possibly $a = 0$. Vice versa, any subset of $\mathbb{Z}$ with the indicated form is a subgroup. So we described all subgroups of $\mathbb{Z}$. ▬

An important property of subgroups of *finite* groups (i.e., groups $(G,\cdot,e)$ such that the set G consists of finitely many elements) is presented in the next theorem. The counting argument used in the proof deserves our attention: we will encounter this technique more often later on.

III.2.7 Theorem. (Theorem of Lagrange; Joseph Louis Lagrange, French mathematician, 1736–1813) *If H is a subgroup of a finite group G, then the number of elements of H is a divisor of the number of elements of G (notation: $\#H|\#G$).*

Proof. For $x \in G$ consider the subset $xH = \{xy \mid y \in H\}$ of G. The union of all subsets of this form is all of G, since $x \in G$ is an element of xH, because $e \in H$ and $x = xe$.

We claim that the subsets considered here have the same number of elements, i.e., $\#xH = \#yH$ for all $x, y \in G$. This follows from the fact that a bijection between xH and yH exists, namely $f : xH \to yH$ given by $f(z) = yx^{-1}z$. Clearly this map sends xH to yH, because if $z \in xH$, then $z = xh$ for some $h \in H$, so $f(z) = yx^{-1}z = yx^{-1}xh = yh \in yH$. The map is bijective, since $g : yH \to xH$ given by $g(z) = xy^{-1}z$ is its inverse, as a short calculation shows. The existence of a bijection between two finite sets means that these sets have the same number of elements.

We will now show that if $xH \cap yH \neq \varnothing$, then the two sets are equal: $xH = yH$. Namely, suppose $z \in xH \cap yH$. Then $z \in xH$, so we may write $z = xh_1$ for some $h_1 \in H$. Similarly $z = yh_2$ for an $h_2 \in H$. Now $xh_1 = yh_2$, and multiplying both sides on the right by h_1^{-1} or by h_2^{-1} shows $x = yh_2h_1^{-1}$ and $y = xh_1h_2^{-1}$. Therefore, we have

written an arbitrary $xh \in xH$ as $xh = yh_2h_1^{-1}h = y(h_2h_1^{-1}h) \in yH$ and similarly an arbitrary $yh \in yH$ as $yh = xh_1h_2^{-1}h \in xH$. This shows $xH = yH$.

We have written G as a union of pairwise disjoint subsets, all having the same number of elements. As a consequence $\#G$ equals the product of the number of such subsets and the cardinality $\#xH = \#eH = \#H$ of the subsets. This implies $\#H|\#G$, which we wanted to prove. ∎

IN order to use this result in the case of subgroups of the form $\langle x \rangle$ we first give the following definition.

III.2.8 Definition. Let x is an element of a group G. Then we define the *order* of x, notation $\mathrm{ord}(x)$, as follows. If $m > 0$ exists with $x^m = e$, then $\mathrm{ord}(x)$ is the least positive integer n such that $x^n = e$. If no such $m > 0$ exists, then $\mathrm{ord}(x) = \infty$.

III.2.9 Example. In every group $(G, \cdot, e)$ it holds that $\mathrm{ord}(e) = 1$. Moreover, if some $x \in G$ satisfies $\mathrm{ord}(x) = 1$ then $x = e$, because $\mathrm{ord}(x) = 1$ implies $x = x^1 = e$.
In $(\mathbb{Z}/5\mathbb{Z})^\times$ we have $\mathrm{ord}(\overline{1}) = 1$ and $\mathrm{ord}(\overline{4}) = 2$ and $\mathrm{ord}(\overline{2}) = \mathrm{ord}(\overline{3}) = 4$.

▬∎

III.2.10 Theorem. *Let G be a group and an element $x \in G$. Then the following statements hold true:*

1. $\mathrm{ord}(x) = \mathrm{ord}(x^{-1})$.
2. *If* $\mathrm{ord}(x) < \infty$, *then* $\langle x \rangle = \{x, x^2, \ldots, x^{\mathrm{ord}(x)} = e\}$.
3. $\mathrm{ord}(x) = \#\langle x \rangle$.
4. *If* $\#G < \infty$, *then also* $\mathrm{ord}(x) < \infty$ *and moreover* $\mathrm{ord}(x)|\#G$.
5. *If* $x^n = e$, *then* $\mathrm{ord}(x)|n$.

Proof. 1: If $x^m = e$, then $(x^{-1})^m = x^{-m} = (x^m)^{-1} = e$. Applying the above for both x and its inverse, we see that the set of integers m with $x^m = e$ equals the set of integers n such that $(x^{-1})^n = e$. (Note that this set may be empty!) In particular it follows that $\mathrm{ord}(x) = \mathrm{ord}(x^{-1})$.
2: Put $d = \mathrm{ord}(x)$. For $m \in \mathbb{Z}$ write $m = qd + r$ with $0 \le r < d$. Then $x^m = (x^d)^q \cdot x^r x^r$. So $\langle x \rangle \subset \{e, x, \ldots, x^{d-1}\}$, which implies the equality.
3: The assertion obviously holds in case $\mathrm{ord}(x) = \infty$ so we will from now on assume that the order of x is finite. In this case 2) implies the result, provided we show that the elements of $\{e, x, \ldots, x^{\mathrm{ord}(x)-1}\}$ are pairwise distinct. Suppose $x^m = x^n$ with $0 \le m \le n < \mathrm{ord}(x)$. Multiplying by the inverse of x^m yields $e = x^{n-m}$, in which $0 \le n - m < \mathrm{ord}(x)$. Since by definition $\mathrm{ord}(x)$ is the least positive d with $x^d = e$, we have $n - m = 0$. So $x^m = x^n$ for non-negative $n, m < ord(x)$ is only possible when $n = m$. This shows 3).
4: $\langle x \rangle$ is a subgroup of G, and since G is finite, so is $\langle x \rangle$. Now 3) and Theorem III.2.7 imply $\mathrm{ord}(x) = \#\langle x \rangle$ is finite and $\mathrm{ord}(x) = \#\langle x \rangle | \#G$.
5: $x^n = e$ implies $\mathrm{ord}(x) < \infty$. Put $d = \gcd(n, \mathrm{ord}(x))$. Then integers k, ℓ exist with $nk + \mathrm{ord}(x)\ell = d$. We have $x^d = (x^n)^k (x^{\mathrm{ord}(x)})^\ell = e$. Since $1 \le d \le \mathrm{ord}(x)$, the definition of $\mathrm{ord}(x)$ implies $d = \mathrm{ord}(x)$. In particular, $\mathrm{ord}(x) = d = \gcd(n, \mathrm{ord}(x))|n$. ∎

III.2.11 Example. In a finite group both the number of elements of any subgroup and the order of any element is a divisor of the number of elements of the group.

However, not every divisor of the number of elements of the group necessarily appears as the order of some element. For example, we already saw that all elements of $(\mathbb{Z}/24\mathbb{Z})^\times$ except the unit element, have order 2. We will return to this issue in a later chapter. ▬∎

III.2.12 Definition. The product of groups
Given two groups $(G_1,\cdot,e_1)$ en $(G_2,*,e_2)$, the product set $G_1\times G_2$ can be given the structure of a group, as follows. This is called the product of groups. By definition, elements of $G_1\times G_2$ are all ordered pairs (x_1,x_2) with $x_i\in G_i$. The unit element in the product group is the pair (e_1,e_2). The group law is given by $(x_1,x_2)\circ(y_1,y_2)=(x_1\cdot y_1,x_2*y_2)$. Check for yourself that indeed with these definitions $(G_1\times G_2,\circ,(e_1,e_2))$ is a group.

Analogously the product of more than two groups is made.

III.2.13 Example. The groups $\mathbb{Z}/8\mathbb{Z}$ and $\mathbb{Z}/4\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$ and $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$ each have 8 elements. Nevertheless in some sense they are very different: the first one contains 4 elements of order 8 while the other two have no such elements. The second group has 4 elements of order 4, whereas the last one only contains elements of order 1 and 2. All three groups are commutative. Incidentally, there are also two "quite different" non-commutative groups with exactly 8 elements (see Exercise 14). ∎

III.3 Homomorphisms

After groups and subgroups we now discuss maps between groups. In courses on *Linear Algebra* the maps considered between vector spaces over a field are the linear maps, i.e. maps preserving the operations (scalar multiplication and addition) defined on vectors. A similar approach will now be done for groups (and various structures within mathematics allow quite analogous treatments of structure preserving maps):

III.3.1 Definition. Given are two groups $(G_1,\cdot,e_1)$ and $(G_2,*,e_2)$. A *homomorphism* from G_1 to G_2 is a map $f:G_1\to G_2$ satisfying $f(x\cdot y)=f(x)*f(y)$ for all $x,y\in G_1$. An *isomorphism* from G_1 to G_2 is a bijective homomorphism.

We call G_1 and G_2 *isomorphic* (notation $G_1\cong G_2$) if an isomorphism from G_1 to G_2 exists.

III.3.2 Example. 1. Let $\mathbb{R}^\times_{>0}$ denote the positive real numbers. This is a group under the usual multiplication. The map $\exp:\mathbb{R}\to\mathbb{R}^\times_{>0}$ given by $x\mapsto e^x$ is an isomorphism from $(\mathbb{R},+,0)$ to $(\mathbb{R}^\times_{>0},\cdot,1)$. Namely, the map exp is bijective, and $\exp(x+y)=e^{x+y}=e^xe^y=\exp(x)\exp(y)$. Hence the two groups are isomorphic.

2. $\det:\mathrm{GL}_n(\mathbb{R})\to\mathbb{R}^\times$ is a homomorphism, since $\det(AB)=\det(A)\det(B)$. For $n\neq 1$ this is not an isomorphism, because for $n>1$ one easily finds two distinct invertible matrices with equal determinant.

3. If G is an arbitrary group and $x\in G$, then $f_x:\mathbb{Z}\to G$ given by $f(n)=x^n$ is a homomorphism. Namely, $f_x(n+m)=x^{n+m}=x^nx^m=f_x(n)f_x(m)$.
In the special case $G=\mathbb{Z}/N\mathbb{Z}$ and $x=\overline{1}$ this map $f_{\overline{1}}$ is the "reduction modulo N": $n\mapsto n \bmod N$.

4. The map $a \bmod NM\mapsto a \bmod N$ used in the Chinese Remainder Theorem II.3.4 is a homomorphism from $(\mathbb{Z}/NM\mathbb{Z},+,0 \bmod NM)$ to $(\mathbb{Z}/N\mathbb{Z},+,0 \bmod N)$ and also from $((\mathbb{Z}/NM\mathbb{Z})^\times,\cdot,1 \bmod NM)$ to $((\mathbb{Z}/N\mathbb{Z})^\times,\cdot,1 \bmod N)$. In particular it follows that if $\gcd(N,M)=1$, then

$$(\mathbb{Z}/NM\mathbb{Z},+,0 \bmod NM)\cong(\mathbb{Z}/N\mathbb{Z},+,0 \bmod N)\times(\mathbb{Z}/M\mathbb{Z},+,0 \bmod M)$$

and

$$((\mathbb{Z}/NM\mathbb{Z})^\times,\cdot,1 \bmod NM)\cong((\mathbb{Z}/N\mathbb{Z})^\times,\cdot,1 \bmod N)\times((\mathbb{Z}/M\mathbb{Z})^\times,\cdot,1 \bmod M).$$

∎

We present some basic properties of homomorphisms.

III.3.3 Theorem. *Given a homomorphism $f : (G_1, \cdot, e_1) \to (G_2, *, e_2)$, the following holds true:*

1. *$f(e_1) = e_2$.*
2. *For all $x \in G_1$ it holds that $f(x^{-1}) = (f(x))^{-1}$.*
3. *If f is an isomorphism, then so is the inverse of f.*
4. *If $g : (G_2, *, e_2) \to (G_3, \star, e_3)$ is a homomorphism as well, then so is the composition $g \circ f$.*

Proof. 1: Write $a = f(e_1)$. In G_2 we have $a * a = f(e_1) * f(e_1) = f(e_1 \cdot e_1) = f(e_1) = a$. Multiplying both sides by the inverse of a yields $a = e_2$.
2: Put $y = f(x^{-1})$. Then $y * f(x) = f(x^{-1}) * f(x) = f(x^{-1} \cdot x) = f(e_1) = e_2$. So y is the inverse of $f(x)$, which is what we wanted to prove.
3: Denote the inverse of the map f by h and let $x, y \in G_2$. Then

$$f(h(x * y)) = x * y = f(h(x)) * f(h(y)) = f(h(x) \cdot h(y)).$$

Since f is bijective this implies $h(x * y) = h(x) \cdot h(y)$. So h is a homomorphism. Bijectivity of f implies that its inverse h is bijective as well. So h is an isomorphism.
4: $g \circ f(x \cdot y) = g(f(x \cdot y)) = g(f(x) * f(y) = g(f(x)) \star g(f(y)) = g \circ f(x) \star g \circ f(y)$ holds for $x, y \in G_1$. ∎

An isomorphism of groups may be considered as a kind of name changer: the elements $x \in G_1$ obtain a new name $f(x) \in G_2$, and the group law is renamed as well. Yet, in a sense nothing has changed. In particular the order of an element remains the same (in spite of the element obtaining a new name). In various cases this observation may be used to show that certain groups are not isomorphic (compare Example III.2.13).

Recall some notations concerning a function φ from a set S_1 to a set S_2: is $T_1 \subset S_1$, then the *image* of T_1 denoted $\varphi(T_1)$, is given by

$$\varphi(T_1) = \{y \in S_2 \mid x \in S_1 \text{ exists with } y = \varphi(x)\}.$$

Similarly for $T_2 \subset S_2$ the *preimage* of T_2 denoted $\varphi^{-1}(T_2)$, is given by

$$\varphi^{-1}(T_2) = \{x \in S_1 \mid \varphi(x) \in T_2\}.$$

The special case of this important concept in group theory is about $\varphi = f$ being a homomorphism of groups, and the T_i being subgroups.

III.3.4 Theorem. *Given is a homomorphism $f : (G_1, \circ, e_1) \to (G_2, *, e_2)$ and subgroups $H_i \subset G_i$ for $i = 1, 2$. Then $f(H_1)$ is a subgroup of G_2, and $f^{-1}(H_2)$ is a subgroup of G_1.*

Proof. For both assertions it suffices to check the conditions H1, H2, and H3. H1: $e_2 \in f(H_1)$, since $e_1 \in H_1$ and $f(e_1) = e_2$. Moreover $e_1 \in f^{-1}(H_2)$, because $f(e_1) = e_2 \in H_2$. Now condition H2: let $x, y \in f^{-1}(H_2)$. Then $f(x), f(y) \in H_2$, so because H_2 is a group, $f(x \cdot y) = f(x) * f(y) \in H_2$ as well. This means $x \cdot y \in f^{-1}(H_2)$. If $w, z \in f(H_1)$, then by definition $u, v \in H_1$ exist with $f(u) = w$ and $f(v) = z$. Now H_1 is a group, so $u \cdot v \in H_1$, and therefore $w * z = f(u) * f(v) = f(u \cdot v) \in f(H_1)$. Finally H3: for $x \in f^{-1}(H_2)$ we know $f(x^{-1}) = (f(x))^{-1} \in H_2$, since $f(x) \in H_2$ and H_2 is a group. This implies $x^{-1} \in f^{-1}(H_2)$. Is $z \in f(H_1)$, then write $z = f(v)$ with $v \in H_1$. Then $z^{-1} = (f(v))^{-1} = f(v^{-1}) \in f(H_1)$, since $v^{-1} \in H_1$. This proves the theorem. ∎

III.3.5 Definition. If $f : (G_1, \cdot, e_1) \to (G_2, *, e_2)$ is a homomorphism, then the *kernel* of f, notation: $\ker(f)$, is defined as

$$\ker(f) = \{x \in G_1 \mid f(x) = e_2\}.$$

III.3.6 Theorem. *Given is a homomorphism $f : (G_1, \circ, e_1) \to (G_2, *, e_2)$.*

1. $\ker(f)$ *is a subgroup of* G_1.
2. f *is injective if and only if* $\ker(f) = \{e_1\}$.

Proof. 1: This is a consequence of Theorem III.3.4, since $\{e_2\}$ is a subgroup of G_2, and $\ker(f) = f^{-1}(\{e_2\})$.
2: We know $f(e_1) = e_2$ by Theorem III.3.3(1.). Let $x \in G_1$ such that $f(x) = e_2 = f(e_1)$. If f is injective, this implies $x = e_1$, so $\ker(f) = \{e_1\}$. Vice versa, let $\ker(f) = \{e_1\}$. If $f(x) = f(y)$ for some $x, y \in G_1$, then $e_2 = f(x) * (f(x))^{-1} = f(y) * f(x^{-1}) = f(y \cdot x^{-1})$. So, $y \cdot x^{-1} \in \ker(f) = \{e_1\}$, i.e., $y \cdot x^{-1} = e_1$. This implies $x = y$, so f is injective. ∎

III.3.7 Remark. Every subgroup of a group G can be written as $f(G')$ for some homomorphism f from some group G' to G. Namely, take the subgroup itself as G', and let f be the inclusion map from G' to G.

However it is *not* possible to realize every subgroup of any group G as the kernel of a homomorphism from G to another group G''. Namely, if $H = \ker(f)$ for a homomorphism f, then H has a property which subgroups in general don't have: is $h \in H$, and is $x \in G$ an arbitrary element, then also $x \cdot h \cdot x^{-1} \in H$. Kernels have this property since $f(x \cdot h \cdot x^{-1}) = f(x) * f(h) * (f(x))^{-1} = f(x) * e_2 * ((f(x))^{-1} = e_2$ for $h \in \ker(f)$. For example the subgroup consisting of all upper triangular matrices in $\mathrm{GL}_2(\mathbb{R})$ does not have the given property. We will see later (Chapter VII) that all subgroups which *do* have the given property, *can* be realised as kernel of some homomorphism.

III.4 Exercises

1. Examine which of the following triples define a group:
   (a) $(\mathbb{N}, +, 0)$;
   (b) $(\mathbb{Q}^\times_{>0}, \cdot, 1)$;
   (c) $(\mathbb{R}, \star, 1)$ with $x \star y = x + y - 1$;
   (d) $(\{x \in \mathbb{R} \mid -\pi/2 < x < \pi/2\}, \circ, 0)$ with $x \circ y = \arctan(\tan(x) + \tan(y))$;
   (e) $(\mathbb{Z}_{>0}, \bullet, 1)$ with $n \bullet m = n^m$.
2. Prove (analogous to the description of all subgroups $\mathbb{Z}$) that the subgroups of $\mathbb{Z}/N\mathbb{Z}$ are exactly all $\langle a \bmod N \rangle$, for $a|N$.
3. Give all subgroups of $(\mathbb{Z}/24\mathbb{Z})^\times$.
4. Consider the 2×2 matrices that (w.r.t. the standard basis of $\mathbb{R}^2$) represent rotation around the origin over 120 degrees and reflection in the x-axis. Construct the smallest possible subgroup of $\mathrm{GL}_2(\mathbb{R})$ containing these two matrices. Is the resulting group abelian? Find the order of each of the elements in this group.
5. Determine all subgroups of the group considered in Exercise 4. Verify for each of them whether it can be written as the kernel of a suitable homomorphism.
6. The *center* $\mathcal{Z}(G)$ of a group G is defined as $\mathcal{Z}(G) = \{x \in G \mid xy = yx \text{ for all } y \in G\}$.
   (a) Show that $\mathcal{Z}(G)$ is an abelian subgroup of G.
   (b) Determine $\mathcal{Z}(\mathrm{GL}_2(\mathbb{R}))$.
7. Given are two finite groups G_1 and G_2. Show that for $(x, y) \in G_1 \times G_2$ it holds that $\mathrm{ord}(x, y) = \mathrm{lcm}(\mathrm{ord}(x), \mathrm{ord}(y))$.
8. In $\mathbb{C}$ we define the subset $\mathbf{T} = \{a + bi \in \mathbb{C} \mid a^2 + b^2 = 1\}$. Denote $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$.
   (a) Show that $(\mathbf{T}, \cdot, 1)$ is a subgroup of $(\mathbb{C}^\times, \cdot, 1)$.
   (b) Prove that $(\mathbb{C}^\times, \cdot, 1) \cong (\mathbf{T}, \cdot, 1) \times (\mathbb{R}_{>0}, \cdot, 1)$.
9. Suppose G is a group and $f : G \to G$ is the map $x \mapsto x \cdot x$. Prove that f is a homomorphism if and only if G is abelian.
10. Suppose $f : G_1 \to G_2$ is a surjective homomorphism of groups. Show that if G_1 is abelian, then so is G_2. Give an example where G_2 is abelian but G_1 is not.
11. (a) Show that an element x in a group G satisfies $x = x^{-1}$ if and only if $\mathrm{ord}(x) = 2$ or $\mathrm{ord}(x) = 1$.
    (b) Conclude that a finite group has an even number of elements if and only if it contains an element of order 2.
12. Show that up to isomorphisms exactly two groups consisting of 4 elements exist (this requires some puzzling; consider the possible orders of elements in such a group, and try to construct the possible multiplication tables).
13. Given a prime p and a group G with exactly p elements. Take $x \in G$ with $x \neq e$. What is the order of x? Prove that $G \cong \mathbb{Z}/p\mathbb{Z}$. So, up to isomorphism only one group consisting of p elements exists (namely, the abelian group $\mathbb{Z}/p\mathbb{Z}$).
14. In the group $\mathrm{GL}_2(\mathbb{C})$ consisting of all invertible 2×2 matrices with complex coefficients we consider two subgroups: H_1 is the minimal one containing both $\left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$; moreover H_2 is the minimal one containing both $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Verify that both subgroups are non-abelian, and that each consists of 8 elements, and that H_1 and H_2 are not isomorphic (for example, count the elements of order 2 in both groups).
15. Let x be an element in a group G, and consider the homomorphism $f_x : \mathbb{Z} \to G$ given by $f(n) = x^n$. Find a relation between the order of x and the kernel of f_x.
16. Given a prime $p \neq 3$ and an $n \in \mathbb{Z}$ with $p|n^2 + n + 1$.
    (a) Verify that $n \bmod p \neq 1 \bmod p$ and that $n \bmod p \in (\mathbb{Z}/p\mathbb{Z})^\times$.

(b) Show that $\operatorname{ord}(n \bmod p) = 3$ in the group $(\mathbb{Z}/p\mathbb{Z})^\times$.

(c) Conclude that $p \equiv 1 \bmod 3$.

(d) Prove that infinitely many primes $\equiv 1 \bmod 3$ exist. (Hint: if $p_1, \ldots, p_t$ are such primes, put $n = 3 \cdot p_1 \cdots p_t$ and consider prime divisors of $n^2 + n + 1$.)

(e) (Compare Exercise 7 in Chapter I): Show that infinitely many primes p exist such that $p + 2$ is not prime.

17. Given is a prime $p \neq 2$ and an $n \in \mathbb{Z}$ with $p | n^2 + 1$.

(a) Verify (analogous to the method used in Exercise 16) that $p \equiv 1 \bmod 4$.

(b) Prove that infinitely many primes $\equiv 1 \bmod 4$ exist.

IV GROUPS OF PERMUTATIONS

In this chapter we study an important class of groups, namely groups consisting of all bijective maps from a set to itself.

IV.1 Bijections of a set

Let Σ be a non-empty set. As we know, a bijection from Σ to itself is a map $\sigma : \Sigma \to \Sigma$ which is injective as well as surjective. Such a σ has a unique inverse, say $\tau : \Sigma \to \Sigma$, satisfying $\sigma \circ \tau = \tau \circ \sigma = \mathrm{id}_\Sigma$. Here, $\circ$ is the composition of maps, and $\mathrm{id}_\Sigma : \Sigma \to \Sigma$ is the identity map given by $\mathrm{id}_\Sigma(x) = x$ for all $x \in \Sigma$. The composition of bijections is a bijection as well.

IV.1.1 Definition. For a non-empty set Σ one denotes by S_Σ the set of all bijections from Σ to itself. The *symmetric group* on the set Σ is by definition the group $(S_\Sigma, \circ, \mathrm{id}_\Sigma)$.

It is easily verified that indeed the symmetric group is a group.

IV.1.2 Example. In case Σ consists of only one element, the only bijection on Σ is the identity. In this case one obtains a group S_Σ consisting of only one element (the "trivial" group).

In case Σ consists of two elements, precisely two bijections are possible: the one fixing both elements (this is id_Σ), and the one interchanging the two elements (let's call it τ). Then $\tau^2 = \tau \circ \tau = \mathrm{id}_\Sigma$. The group S_Σ is in this case isomorphic to $\mathbb{Z}/2\mathbb{Z}$.

For Σ with $\#\Sigma > 2$ the group S_Σ is not commutative. Namely, take three distinct elements $x, y, z \in \Sigma$. Define two bijections $\sigma, \tau \in S_\Sigma$ as follows. σ interchanges x and y and it fixes all other elements of Σ. Similarly τ interchanges y and z and it fixes the remaining elements. This indeed defines two bijections, and $\sigma \circ \tau(x) = y$ whereas $\tau \circ \sigma(x) = z$. So $\sigma \circ \tau \neq \tau \circ \sigma$. In particular, S_Σ is not commutative. ∎

Given two "equally big" sets Σ and Σ' (more precisely: two sets with a bijection $f : \Sigma \xrightarrow{\sim} \Sigma'$), then intuitively it may be clear that the groups S_Σ and $S_{\Sigma'}$ are isomorphic. Indeed, the bijection f provides a way to give all elements of Σ a new name, and describing bijections in terms of either the old or the new names is essentially the same. Turning this argument into a formal proof yields the following:

IV.1.3 Theorem. *Suppose $f : \Sigma \to \Sigma'$ is a bijection and $g : \Sigma' \to \Sigma$ is its inverse (so $f \circ g = \mathrm{id}_{\Sigma'}$ and $g \circ f = \mathrm{id}_\Sigma$). Then S_Σ and $S_{\Sigma'}$ are isomorphic; an explicit isomorphism $\varphi : S_\Sigma \to S_{\Sigma'}$ is given by $\varphi(\sigma) = f \circ \sigma \circ g$, with as inverse $\psi : S_{\Sigma'} \to S_\Sigma$ given by $\psi(\tau) = g \circ \tau \circ f$.*

Proof. This is a useful exercise in formal calculations with compositions of maps, and it tests understanding of a number of definitions. We leave it as an exercise! ∎

A special case is obtained by considering only finite sets Σ. A bijection between two such sets exists precisely when they have the same number of elements. Hence the result above shows that when studying symmetry groups of finite sets, it suffices to consider the sets $S_{\{1,2,\ldots,n\}}$. We present another example of an abstract theorem concerning groups of bijections of a set.

IV.1.4 Theorem. (Cayley's theorem; Arthur Cayley, English mathematician, 1821–1895) *Every group G is isomorphic to a subgroup of S_G. In the special case that $\#G = n < \infty$, the group G is isomorphic to a subgroup of $S_{\{1,\ldots,n\}}$.*

Proof. For fixed $a \in G$ the map $\lambda_a : G \to G$ given by $\lambda_a(x) = ax$ is a bijection (see Theorem III.1.5). So $\lambda_a \in S_G$. This is used to define a map

$$\varphi : G \longrightarrow S_G$$

by $\varphi(a) = \lambda_a$. One easily verifies that φ is a homomorphism, i.e., for $a, b \in G$ one has $\varphi(ab) = \lambda_{ab} = \lambda_a \circ \lambda_b = \varphi(a) \circ \varphi(b)$.

The homomorphism φ is injective because any $a \in \ker(\varphi)$ satisfies by definition $\lambda_a = \mathrm{id}_G$, so $a = ae = \lambda_a(e) = \mathrm{id}_G(e) = e$. It follows that G is isomorphic to $\varphi(G)$, and the latter is indeed a subgroup of S_G.

The assertion concerning the case $\#G = n$ follows by combining the above with Theorem IV.1.3. ∎

IV.2 Permutations on n integers

IV.2.1 Definition. Let $n \in \mathbb{Z}_{\geq 1}$. The *symmetric group on n integers*, notation S_n, is by definition the group $S_{\{1,2,\ldots,n\}}$. Elements of this group are called *permutations*. S_n is also called the permutation group on n elements.

IV.2.2 Theorem. *The group S_n consists of $n! = n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1$ elements.*

Proof. An element of S_n is by definition a bijection of the set $\{1, \ldots, n\}$. Such a bijection can be described as a sequence of length n, in which each of $1, \ldots, n$ appears exactly once. One readily verifies that precisely $n!$ such sequences exist. ∎

IV.2.3 Definition. A permutation $\sigma \in S_n$ is called a *cycle* of length k (or, k-cycle), in case k distinct integers $a_1, \ldots, a_k \in \{1, \ldots, n\}$ exist with $\sigma(a_i) = a_{i+1}$ for $1 \leq i < k$ and $\sigma(a_k) = a_1$ and $\sigma(x) = x$ for $x \notin \{a_1, \ldots, a_k\}$. This is denoted by $\sigma = (a_1\, a_2\, \ldots\, a_k)$. A 2-cycle is also called a transposition.
When two cycles $(a_1\, a_2\, \ldots\, a_k)$ and $(b_1\, b_2\, \ldots\, b_\ell)$ satisfy $\{a_1, \ldots, a_k\} \cap \{b_1, \ldots, b_\ell\} = \varnothing$, they are called *disjoint*.

IV.2.4 Example. We have $(1\,2\,3\,4\,5) = (2\,3\,4\,5\,1) = \ldots = (5\,1\,2\,3\,4)$, since these 5-cycles send 5 to 1, and i to $i+1$ for $1 \leq i < 5$, and they fix the integers ≥ 6. In general, the same reasoning shows for k-cycles that $(a_1\, a_2\, \ldots\, a_k) = (a_2\, \ldots\, a_k\, a_1)$.

Two disjoint cycles commute: namely, if $(a_1\, a_2\, \ldots\, a_k)$ and $(b_1\, b_2\, \ldots\, b_\ell)$ are disjoint, then the first one only affects the integers $a_1, \ldots, a_k$ and the second one only $b_1, \ldots, b_\ell$. Hence it is irrelevant in which order these cycles are applied.

This is different for non-disjoint cycles: e.g., $(1\,2\,3) \circ (2\,3\,4) \neq (2\,3\,4) \circ (1\,2\,3)$, since the first composition maps 2 to 1 and the second one maps 2 to 4. (Note that we are composing functions, so the rightmost function is applied first!) ∎

IV.2.5 Theorem. *Every $\sigma \in S_n$ can be written as a product $\sigma = \sigma_1 \dots \sigma_r$, with the σ_i pairwise disjoint cycles. Apart from the order of the σ_i, this presentation is unique.*

Proof. The existence can be shown using induction w.r.t. n, as follows. For $n = 1$ the assertion is clear, namely in this case the only permutation is $\sigma = (1)$. Let $n > 1$ and assume the existence for all S_m with $m < n$. Let $\sigma \in S_n$, then $\{1, \sigma(1), \sigma^2(1), \dots\}$ is a subset of $\{1, \dots, n\}$, so k, ℓ exist with $k < \ell$ and $\sigma^k(1) = \sigma^\ell(1)$. One concludes $\sigma^{\ell-k}(1) = 1$, so a positive integer s exists with $\sigma^s(1) = 1$. The least such integer we denote by q. By construction the integers $1, \sigma(1), \dots, \sigma^{q-1}(1)$ are pairwise distinct, and the effect of σ on these integers is given by the k-cycle $\sigma_1 = (1\ \sigma(1)\ \dots\ \sigma^{q-1}(1))$.

Now consider the remaining integers in $\{1, \dots, n\}$. In case this is the empty set, then $\sigma = \sigma_1$ and we are done. If the set is nonempty, then σ acts as a permutation on it. The induction hypothesis implies that the restriction of σ to this subset can be written as a product of disjoint cycles $\sigma_2 \dots \sigma_r$. Considering these cycles as permutations on $\{1, \dots, n\}$, we have $\sigma = \sigma_1 \dots \sigma_r$.

To show uniqueness, imagine that some permutation allows two different presentations as product of disjoint cycles. Suppose $i \mapsto j$, then in both of the presentations exactly one cycle occurs containing $(\dots i\ j \dots)$. Now considering the image of j etc., shows that the presentations contain the same cycles, so they are equal. ∎

IV.2.6 Example. The argument above is in fact an algorithm. To illustrate this, suppose we want to write $(1\,2\,3\,4)(2\,3\,4\,5)(4\,5\,1)$ as a product of disjoint cycles. We see here a composition of maps. First, we determine its effect on 1. The rightmost permutation sends 1 to 4, and 4 is mapped by the middle permutation to op 5. The leftmost permutation fixes 5, so in total the image of 1 is 5. Next we find out what happens to 5. The rightmost sends 5 to 1; this 1 is fixed by the middle one and than sent to 2 by the remaining cycle. Continuing in this way we find that 4 is the image of 2, and 4 is mapped to 3, and 3 to 1. In this way we have found a 5-cycle, and since the initial permutations only involve the integers 1 to 5, we are done: $(1\,2\,3\,4)(2\,3\,4\,5)(4\,5\,1) = (1\,5\,2\,4\,3)$. ▬∎

Writing a permutation as a product of disjoint cycles helps us to determine the order of a permutation:

IV.2.7 Theorem. *1.* $(i_1\ i_2\ \dots\ i_k)^{-1} = (i_k\ i_{k-1}\ \dots i_1)$.
2. A k-cycle $(i_1\ i_2\ \dots\ i_k)$ has order k.
3. Are $\sigma_1, \dots, \sigma_r$ pairwise disjoint cycles, then $(\sigma_1 \dots \sigma_r)^n = \sigma_1^n \dots \sigma_r^n$ for all $n \in \mathbb{Z}$.
4. If moreover σ_i has length ℓ_i $(i = 1, \dots, r)$, then $\mathrm{ord}(\sigma_1 \dots \sigma_r) = \mathrm{lcm}(\ell_1, \dots, \ell_r)$.

Proof. 1: This is immediate from the definition of a cycle.
2: For $0 < n < k$ the image of i_1 under $(i_1\ i_2\ \dots\ i_k)^n$ equals i_{n+1}. Since $i_{n+1} \neq i_1$, this means the cycle has order $\geq k$. Now $(i_1\ i_2\ \dots\ i_k)^k = (1)$, so the order equals k.
3: This is a consequence of the fact that disjoint cycles commute.
4: Using 3) and the uniqueness in Theorem IV.2.5 it follows that $(\sigma_1 \dots \sigma_r)^n = (1)$ precisely when $\sigma_1^n = \dots = \sigma_r^n = (1)$. By Theorem III.2.10 the latter holds if and only if n is a multiple of each of $\mathrm{ord}(\sigma_1), \dots, \mathrm{ord}(\sigma_r)$. ∎

IV.2.8 Example. The n-th power of a k-cycle, with $1 < n < k$, is not necessarily itself a k-cycle. As an example, $(1\,2\,3\,4)^2 = (1\,3)(2\,4)$. ▬∎

IV.2.9 Example. We determine which integers occur as order of some element in S_5. Note $5 = 4+1 = 3+2 = 3+1+1 = 2+2+1 = 2+1+1+1 = 1+1+1+1+1$. These are all presentations of 5 as a sum of positive integers. It shows that a product of disjoint cycles in S_5 can be obtained in 7 ways: a 5-cycle, or a 4-cycle (multiplied by a 1-cycle, which we leave out since it represents the identity map), or

… etc. Theorem IV.2.7 implies that the orders of these products are $5, 4, 6, 3, 2$, and 1, respectively. For each of these numbers it is a relatively simple combinatorial problem to determine how many elements in S_5 have as order the given number. ▬■

IV.2.10 Theorem. *Every $\sigma \in S_n$ can be written as a product of 2-cycles.*

Proof. We know that σ is a product of cycles. So it suffices to write any cycle as a product of 2-cycles:

$$(a_1\, a_2\, \ldots\, a_k) = (a_1\, a_2)(a_2\, a_3)\ldots(a_{k-1}\, a_k)$$

as is readily checked. ■

IV.2.11 Remark. Theorem IV.2.10 states in fact that by repeatedly interchanging pairs (i.e., performing transpositions), a row of n objects can be placed in an arbitrary order. The proof even provides an upper bound for the number of transpositions needed: given a permutation as product of r disjoint ℓ_i-cycles, with $\ell_i \geq 1$ and $\sum \ell_i = n$. For an ℓ_i-cycle the proof of Theorem IV.2.10 shows that $\ell_i - 1$ interchanges suffice. In total we therefore obtain the upper bound $\sum(\ell_i - 1) = n - r$.

IV.2.12 Remark. It is even possible to show that any permutation can be written as a product of transpositions of a special kind. For example, using only 2-cycles $(i\ i+1)$: If $i < j$, then

$$(i\ j) = (i\ i+1)(i+1\ i+2)\ldots(j-1\ j)(j-2\ j-1)\ldots(i\ i+1).$$

Alternatively, one can write any permutation as a product of 2-cycles $(1\ i)$. This follows from the observation that for $1 \neq i \neq j \neq 1$ one has $(i\ j) = (1\ i)(1\ j)(1\ i)$. In other words: by merely interchanging one given element consecutively with suitable other(s), a row of objects can be put in arbitrary order.

Note that the presentation of a permutation as product of 2-cycles is far from unique! In the next section we will show that the *parity* of the number of 2-cycles needed to represent a given permutation, is fixed.

IV.3 Even and odd permutations

IV.3.1 Notation. 1. For $n \geq 2$ write $X := \{(i,j) \in \mathbb{Z} \times \mathbb{Z} \mid 1 \leq i < j \leq n\}$.
2. For $\sigma \in S_n$ define $f_\sigma : X \to X$ by $f_\sigma(i,j) = (\min\{(\sigma(i), \sigma(j)\}, \max\{(\sigma(i), \sigma(j)\})$.
3. Finally, define $h_\sigma : X \to \mathbb{Q}$ by $h_\sigma(i,j) = \dfrac{\sigma(j) - \sigma(i)}{j - i}$.

Some useful properties of these functions are as follows.

IV.3.2 Lemma. *As before, let $n \geq 2$.*

1. *For $\sigma, \tau \in S_n$ one has $f_{\sigma\tau} = f_\sigma \circ f_\tau$.*
2. *f_σ is a bijection on X.*
3. $\displaystyle\prod_{(i,j)\in X} h_\sigma(i,j) = \pm 1.$

Proof. 1: Both functions map an arbitrary $(i,j) \in X$ to either $(\sigma\tau(i), \sigma\tau(j))$, or to $(\sigma\tau(j), \sigma\tau(i))$ (depending on which of the two is in X). So the functions coincide.
2: This follows from $f_\sigma \circ f_{\sigma^{-1}} = f_{\sigma^{-1}} \circ f_\sigma = f_{id} = \mathrm{id}$.
3: The absolute value of the given product equals

$$\left(\prod_{(i,j)\in X} |\sigma(j) - \sigma(i)| \right) / \left(\prod_{(i,j)\in X} (j-i) \right).$$

Here the numerator is the product of all $(\ell - k)$, for $(k,\ell) = f_\sigma(i,j) \in f_\sigma(X) = X$. So numerator and denominator are equal. The absolute value being 1 implies that the product is ± 1. ∎

IV.3.3 Definition. The *sign* of a permutation $\sigma \in S_n$, notation $\epsilon(\sigma)$, is by definition

$$\epsilon(\sigma) = \prod_{(i,j)\in X} h_\sigma(i,j) = \prod_{1\le i<j\le n)} \frac{\sigma(j) - \sigma(i)}{j-i} = \pm 1$$

in case $n \ge 2$, and $\epsilon(\sigma) = 1$ for $n = 1$. We call σ *even* if $\epsilon(\sigma) = 1$ and *odd* if $\epsilon(\sigma) = -1$.

IV.3.4 Remark. We will soon describe an efficient way to calculate the sign of a permutation. Only using the definition, this may be quite elaborate; as an example, try to determine the sign of the 3-cycles (1 3 5) and (1 6 12)!

The sign of a permutation may be interpreted as follows: the denominator of the expression defining the sign is a product of positive integers. The factors of the numerator have the form $\sigma(j) - \sigma(i)$, and this factor is negative precisely when σ maps the integers i and j to a pair which is, w.r.t. the usual 'smaller than'-relation on $\{1,\dots,n\}$ in the other order as i,j. If this occurs for an *even* number of pairs $(i,j) \in X$, then the sign $\epsilon(\sigma) = 1$; if it happens for an *odd* number of pairs, then σ has sign -1.

The set $\{+1,-1\}$ is a group w.r.t. the usual multiplication. So ϵ is a map from the group S_n to the group ± 1.

IV.3.5 Theorem. *The sign $\epsilon : S_n \longrightarrow \pm 1$ is a homomorphism.*

Proof. Let $\sigma, \tau \in S_n$. Then

$$\prod_{1\le i<j\le n} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} = \prod_{(i,j)\in X} h_\sigma(f_\tau(i,j)) = \prod_{(i,j)\in X} h_\sigma(i,j) = \epsilon(\sigma),$$

since f_σ is bijective on X (Lemma IV.3.2). Hence

$$\begin{aligned} \epsilon(\sigma\tau) &= \textstyle\prod_{(i,j)\in X} \frac{(\sigma\tau)(j) - (\sigma\tau)(i)}{j-i} \\ &= \left(\textstyle\prod_{1\le i<j\le n} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \right) \left(\textstyle\prod_{(i,j)\in X} \frac{\tau(j) - \tau(i)}{j-i} \right) \\ &= \epsilon(\sigma)\epsilon(\tau). \end{aligned}$$

This proves the theorem. ∎

In order to be able to use this result for computing the sign of permutations, we first prove a lemma. The first part of the lemma will also be important later on for showing more properties of permutations.

IV.3.6 Lemma. 1. $\tau(a_1\, a_2\, \dots\, a_\ell)\tau^{-1} = (\tau(a_1)\, \tau(a_2)\, \dots\, \tau(a_\ell))$ for any $\tau \in S_n$ and any ℓ-cycle $(a_1\, a_2\, \dots\, a_\ell) \in S_n$.
2. A 2-cycle $(a_1\, a_2)$ satisfies $\epsilon((a_1\, a_2)) = -1$.

Proof. 1: We have $(\tau(a_1\ \ldots\ a_\ell)\tau^{-1})(\tau(a_\ell)) = (\tau(a_1\ \ldots\ a_\ell))(a_\ell) = \tau(a_1)$. Similarly for $1 \le k < \ell$ one finds $(\tau(a_1\ a_2\ \ldots\ a_\ell)\tau^{-1})(\tau(a_k)) = \tau(a_{k+1})$. For all remaining $i \in \{1,\ldots,n\}$ one has $(\tau(a_1\ a_2\ \ldots\ a_\ell)\tau^{-1})(i) = i$. This shows the equality.
2: Take any permutation τ with $\tau(a_1) = 1$ and $\tau(a_2) = 2$. Then $\epsilon((1\,2)) = \epsilon(\tau(a_1\ a_2)\tau^{-1})$. Since ϵ is a homomorphism, we have $\epsilon(\tau(a_1\ a_2)\tau^{-1}) = \epsilon(\tau)\epsilon((a_1\ a_2))\epsilon(\tau)^{-1} = \epsilon((a_1\ a_2))$. So all 2-cycles have the same sign. For $(1\,2)$ we determine the sign from the definition: $\epsilon((1\,2)) = -1$, because the only pair $(i,j) \in X$ changing order when $(1\,2)$ is applied, is $(1,2)$. ∎

IV.3.7 Corollary. *1. An ℓ-cycle σ has sign $\epsilon(\sigma) = (-1)^{\ell-1}$.*
2. If σ is a product of cycles of lengths $\ell_1,\ldots,\ell_r$, then $\epsilon(\sigma) = (-1)^{\sum_{i=1}^{r}(\ell_i-1)}$.
3. A permutation σ is even if and only if σ can be written as a product of an even number of 2-cycles.

Proof. 1: We have $(a_1\ a_2\ \ldots\ a_\ell) = (a_1\ a_2)(a_2\ a_3)\ldots(a_{\ell-1}\ a_\ell)$. The number of 2-cycles in this product is $\ell - 1$, so because all of them have sign -1 and ϵ is a homomorphism, the result follows.
2: This is immediate from 1) since ϵ is a homomorphism.
3: Let $\sigma \in S_n$. By Theorem IV.2.10 σ can be written as a product of 2-cycles. The assertion now follows from 2) above. ∎

IV.4 The alternating group

IV.4.1 Definition. For $n \ge 1$ the *alternating group* (notation: A_n) is the subgroup of S_n consisting of all even permutations.

The fact that indeed A_n is a group follows from $A_n = \ker(\epsilon)$ and ϵ is a homomorphism (Theorems IV.3.5 and III.3.6).

IV.4.2 Example. S_2 consists of the permutations (1) and $(1\,2)$. So $A_2 = \{(1)\}$.

S_3 consists of the identity, 2-cycles, and 3-cycles. The 2-cycles are not in A_3, and $A_3 = \{(1),(1\,2\,3),(1\,3\,2)\}$. This group is isomorphic to $\mathbb{Z}/3\mathbb{Z}$.

In A_4 one finds apart from the identity the 3-cycles (there are 8 of them) and the products of two disjoint 2-cycles (in total 3). The group obtained in this way is not abelian, and consists of 12 elements. ∎

IV.4.3 Theorem. *For $n \ge 2$ the group A_n consists of $n!/2$ elements.*

Proof. The sets A_n and $(S_n \setminus A_n)$ are by definition disjoint, with union all of S_n. They have the same number of elements, because the map $\tau \mapsto (1\,2)\tau$ is a bijection between them. The result follows. ∎

IV.4.4 Theorem. For $n \ge 3$ the elements of A_n can be written as products of 3-cycles.

Proof. Let $\sigma \in A_n$. By Corollary IV.3.7 σ is a product of an even number of 2-cycles. In particular σ is a product of permutations $(a\ b)(c\ d)$. If $\{a,b\} = \{c,d\}$ the latter equals (1). If $\{a,b\}$ and $\{c,d\}$ have one element, say $a = c$ in common, then $(a\ b)(a\ d) = (a\ d\ b)$ is a 3-cycle. In the remaining case $(a\ b)(c\ d) = (a\ c\ b)(c\ d\ a)$. This shows the theorem. ∎

IV.4.5 Example. We finish this chapter by illustrating Cayley's theorem presented at the beginning (Theorem IV.1.4). Take $G = (\mathbb{Z}/8\mathbb{Z})^*$. Since $\#G = \varphi(8) = 4$, the group G is isomorphic to a subgroup of S_4. We determine which subgroup the

proof of Theorem IV.1.4 yields, and we even show that this subgroup is contained in A_4. The presented proof identifies $a \in G$ with λ_a, the left-multiplication by a map. Moreover S_G is identified with S_4, simply by choosing a bijection between G and $\{1,2,3,4\}$. We choose the bijection $\overline{1} \mapsto 1, \overline{3} \mapsto 2, \overline{5} \mapsto 3$ en $\overline{7} \mapsto 4$.

The element $\overline{1} \in G$ gives rise to $\lambda_{\overline{1}} = \mathrm{id}_G$, which is the permutation (1). The element $\overline{3}$ yields the bijection on G sending $\overline{1}$ to $\overline{3}$, $\overline{3}$ to $\overline{3}\cdot\overline{3} = \overline{1}$, $\overline{5}$ to $\overline{3}\cdot\overline{5} = \overline{7}$, and $\overline{7}$ to $\overline{5}$. Using our bijection between G and $\{1,2,3,4\}$ this becomes the permutation (1 2)(3 4).

A similar calculation sends $\overline{5}$ to the permutation (1 3)(2 4) and $\overline{7}$ to (1 4)(2 3). So apparently $\{(1),(1\,2)(3\,4),(1\,3)(2\,4),(1\,4)(2\,3)\}$ is a subgroup of S_4, isomorphic to $(\mathbb{Z}/8\mathbb{Z})^*$. Clearly this subgroup is contained in A_4. ∎

IV.5 Exercises

1. Write each of the following permutations as product of disjoint cycles:
   (a) $(3\,1\,4)(1\,5\,9\,2\,6)(5\,3)$
   (b) σ^{-1}, for $\sigma = (5\,6\,2)(1\,3)(1\,4)$.
2. (a) Find all $\sigma \in S_4$ satisfying $\sigma^2 = (1\,2)(3\,4)$.
   (b) Let $n > 1$. Does $\sigma \in S_n$ exist such that $\sigma^2 = (1\,2)$?
   (c) Let $n \geq 6$. Does $\sigma \in S_n$ exist such that $\sigma^2 = (1\,2)(3\,4\,5\,6)$?
3. Suppose σ is a k-cycle is. Show: σ^n is a k-cycle if and only if $\gcd(k,n) = 1$.
4. (a) Determine which integers occur as order of an element of S_6.
   (b) For each of the integers above, how many elements in S_6 have this order?
5. What is the least n such that $30 | \#S_n$? What is the least n such that S_n contains an element of order 30?
6. Determine the order and the sign of $(5\,6\,7\,8\,9)(3\,4\,5\,6)(2\,3\,4)(1\,2)$ in S_9.
7. (a) With $\sigma = (1\,2)(3\,4\,5)(6\,7\,8\,9\,10)$, write σ^{2016} as a product of disjoint cycles.
   (b) Do the same with τ^{2017}, given $\tau = (1\,2\,3)(3\,4)(4\,5\,6\,7)$.
8. Let $\sigma \in S_n$. Show that if $\sigma(1\,2\,\ldots\,n) = (1\,2\,\ldots\,n)\sigma$, then $\sigma = (1\,2\,\ldots\,n)^i$ some i.
9. For $n \geq 1$, determine the center $\mathcal{Z}(S_n)$ of S_n (these are the $\tau \in S_n$ satisfying $\sigma\tau = \tau\sigma$ for all $\sigma \in S_n$).
10. (a) Let $\sigma, \tau \in S_n$. Show that if σ is a product of disjoint cycles of lengths $\ell_1, \ldots, \ell_r$, then so is $\tau\sigma\tau^{-1}$.
    (b) Vice versa, if $\sigma_1, \sigma_2 \in S_n$ are both products of disjoint cycles of lengths $\ell_1, \ldots, \ell_r$, show that $\tau \in S_n$ exists with $\sigma_2 = \tau\sigma_1\tau^{-1}$.
11. (a) Suppose $a \neq 1 \neq b$. Compute $(1\,a)(1\,b)(1\,a)(1\,b)$.
    (b) Show that every element of A_n can be written as a product of elements of the form $\sigma\tau\sigma^{-1}\tau^{-1}$, for $\sigma, \tau \in S_n$.
    (c) Show that if G is an abelian group and $f : S_n \to G$ a homomorphism, then $A_n \subset \ker(f)$.
    (d) Show that if $g : S_n \to S_m$ is a homomorphism, then $g(A_n) \subset A_m$.
12. A subgroup $H \subset S_n$ is called *transitive* if for every $\{i, j\} \subset \{1, 2, \ldots, n\}$ some $\tau \in H$ exists with $\tau(i) = j$.
    (a) Show that for $n \geq 3$ the group A_n is a transitive subgroup of S_n.
    (b) Show that if G is a group and $\#G = n$, then the subgroup of S_n constructed in the proof of Cayley's theorem is a transitive subgroup of S_n.
    (c) Using Cayley's theorem, construct a transitive subgroup of S_6 isomorphic to S_3.

V GROUPS OF SYMMETRIES

In this chapter groups consisting of special bijections on some space or set, are considered. This leads to the kind of groups which in particular are of interest to physics, or in some cases also to discrete mathematics. The concepts from Linear Algebra which are used in the chapter may be found in essentially all textbooks on the subject.

V.1 Some groups of matrices

The vector space $\mathbb{R}^2$ over $\mathbb{R}$ can be visualized as a plane. The standard way to do so we learn already in the first years of high school. Using the theorem of Pythagoras, this standard interpretation allows us to introduce a distance function d on $\mathbb{R}^2$:

$$d((a,b),(c,d)) = \sqrt{(a-c)^2+(b-d)^2}.$$

We learned an analogous distance for $\mathbb{R}^3$, and in Linear Algebra this has been generalized to the situation of an arbitrary (real or complex hermitian) inner product space $(V,\langle\cdot\,,\cdot\rangle)$. In the latter case the distance $d(v,w)$ between two vectors $v,w\in V$ is defined as

$$d(v,w) = ||v-w|| = \sqrt{\langle v-w, v-w\rangle}.$$

Attached to such an inner product space is a group:

V.1.1 Definition. Let $(V,\langle\cdot\,,\cdot\rangle)$ be a real or complex hermitian inner product space. The set of all linear maps $\varphi : V\to V$ satisfying $\langle v,w\rangle = \langle\varphi(v),\varphi(w)\rangle$ for all $v,w\in V$, is denoted as $O(V,\langle\cdot\,,\cdot\rangle)$.

V.1.2 Theorem. *For $(V,\langle\cdot\,,\cdot\rangle)$ as above and moreover V finite dimensional, the set $O(V,\langle\cdot\,,\cdot\rangle)$ is w.r.t. the composition of linear maps a group, with unit element* id_V.

Proof. We first show that $O(V,\langle\cdot\,,\cdot\rangle)$ is a subset of the group $\mathrm{GL}(V)$ consisting of *all* invertible linear maps from V to itself. In other words, we show that the elements of $O(V,\langle\cdot\,,\cdot\rangle)$ are invertible. Let $\varphi\in O(V,\langle\cdot\,,\cdot\rangle)$. If $\varphi(v)=0$, then $\langle v,v\rangle = \langle\varphi(v),\varphi(v)\rangle = 0$, hence $v=0$. This implies that φ is injective. Since injective linear maps from a finite dimensional vector space to itself are automatically surjective, φ is invertible.

Hence to show that $O(V,\langle\cdot\,,\cdot\rangle)$ is a group, it suffices to verify that it is a subgroup of $\mathrm{GL}(V)$. With the aid of Theorem III.2.3 we leave this as an exercise to the reader. ∎

V.1.3 Remark. Note that the condition "V is finite dimensional" was used in the argument above to show that indeed elements of $O(V,\langle\,\cdot\,,\,\cdot\,\rangle)$ are invertible. In fact, for inner product spaces of infinite dimension over $\mathbb{R}$ or $\mathbb{C}$ this may not hold, as the following example shows. Take V the vector space over $\mathbb{R}$ consisting of all sequences $(a_n)_{n\geq 1}$ of real numbers, with the property that $a_n = 0$ for all but finitely many positive integers n. On V we define the 'standard' inner product $\langle (a_n)_{n\geq 1},(b_n)_{n\geq 1}\rangle = \sum_{n=1}^{\infty} a_n b_n$. (The sum is well defined since only finitely many terms are nonzero.) The shift operator $\sigma : V \to V$ given by

$$\sigma(a_1,a_2,a_3,\ldots) = (0,a_1,a_2,a_3,\ldots)$$

is an element of $O(V,\langle\,\cdot\,,\,\cdot\,\rangle)$, but it is clearly not invertible. So in this example $O(V,\langle\,\cdot\,,\,\cdot\,\rangle)$ is *not* a group.

If $A = (a_{i,j})$ is a (square) matrix with real or complex coefficients, then Linear Algebra defines the adjoint of A, notation A^*, as $A^* = (b_{i,j})$ with $b_{i,j} = \overline{a_{j,i}}$. So to obtain A^* one reflects all entries of the matrix w.r.t. the main diagonal, and then one takes the complex conjugate of the entries. A more intrinsic definition of A^* is that w.r.t. the standard inner product on $\mathbb{R}^n$ or $\mathbb{C}^n$, it is the unique matrix such that for all vectors v,w one has $\langle Av,w\rangle = \langle v,A^*w\rangle$. More generally, is $\varphi \in \mathrm{GL}(V)$ w.r.t. an orthonormal basis of V given by a matrix A, then $\varphi \in O(V,\langle\,\cdot\,,\,\cdot\,\rangle)$ if and only if $A^*A = I$, with I the unit matrix. This translates the group $O(V,\langle\,\cdot\,,\,\cdot\,\rangle)$ into a group of matrices, namely into a subgroup of $\mathrm{GL}_n(\mathbb{R})$ of $\mathrm{GL}_n(\mathbb{C})$ with $n = \dim(V)$. We now mention some groups of matrices and some relevant subgroups obtained in this way.

V.1.4 Definition. Let $n \in \mathbb{Z}, n > 0$.

1. The *orthogonal* group $\mathrm{O}(n) = \{A \in \mathrm{GL}_n(\mathbb{R}) \mid A^*A = I\}$.
2. The *unitary* group $\mathrm{U}(n) = \{A \in \mathrm{GL}_n(\mathbb{C}) \mid A^*A = I\}$.
3. The *special orthogonal* group $\mathrm{SO}(n) = \{A \in \mathrm{GL}_n(\mathbb{R}) \mid A^*A = I \text{ and } \det(A) = 1\}$.
4. The *special unitary* group $\mathrm{SU}(n) = \{A \in \mathrm{GL}_n(\mathbb{C}) \mid A^*A = I \text{ and } \det(A) = 1\}$.

V.1.5 Example. For $n = 1$ we obtain $\mathrm{O}(1) = \{a \in \mathbb{R}\setminus\{0\} \mid a^2 = 1\} = \{\pm 1\}$. As group of maps on $\mathbb{R}$ these are the identity and 'taking the opposite'. The groups SO(1) and SU(1) both equal the trivial group consisting of only one element. U(1) is more interesting: it is the group of all points on the unit circle in $\mathbb{C}$, with multiplication as group law.

The group SO(2) consists of all matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $a,b,c,d \in \mathbb{R}$ satisfying $a^2 + c^2 = b^2 + d^2 = ad - bc = 1$ and $ab + cd = 0$. Writing $a = \cos\alpha$ and $c = \sin\alpha$, it follows that $d = \cos\alpha$ and $b = -\sin\alpha$. So as a map from $\mathbb{R}^2$ to itself this matrix represents the rotation with center $(0,0)$ over an angle α. The group SO(2) is exactly the group consisting of all such rotations.

In the group O(2) we find apart from the matrices in SO(2) also those given as $\begin{pmatrix} \cos\alpha & -\sin\alpha \\ -\sin\alpha & -\cos\alpha \end{pmatrix}$. Geometrically this represents a reflection in the line passing through the origin, which intersects the positive x-as with an angle $-\alpha/2$. So O(2) consists of geometrically defined maps, namely all rotations with the origin as center, and all refections in lines passing through the origin. O(2) is not commutative: if we first reflect in the x-axis and then rotate (counter clockwise) over an angle of 90 degrees, $(0,1)$ is the image of $(1,0)$. However applying the maps in the reverse order maps $(1,0)$ to $(0,-1)$. ∎

All groups given in Definition V.1.4 may be regarded as groups of invertible linear maps from $\mathbb{R}^n$ or $\mathbb{C}^n$ to itself, with the property that they preserve the standard inner product and therefore also the distance between points.

From now on we will restrict ourselves to the real case, and in particular to $\mathbb{R}^2$ and $\mathbb{R}^3$. Geometrically the fact that a map is distance preserving means, that for example a triangle is mapped to a congruent triangle. Namely, the three vertices are mapped to three new points which pairwise have the same distance as the original ones. A point on one of the sides is mapped to a point which has the same distances to the new vertices as the original point had to the old ones. This implies that the sides of the original triangle are mapped to sides of the new one. This argument shows that all distance preserving maps (we do not need to assume linearity) map lines to lines and angles to equally large angles.

V.2 Groups of isometries

We work with the space $\mathbb{R}^n$, equipped with the norm $||(a_1,\ldots,a_n)|| = \sqrt{a_1^2+\ldots a_n^2}$ and the distance $d(v,w) = ||v-w||$ for $v,w \in \mathbb{R}^n$.

V.2.1 Definition. An *isometry* on $\mathbb{R}^n$ is a map $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ with the property $d(v,w) = d(\varphi(v),\varphi(w))$ for all $v,w \in \mathbb{R}^n$.

V.2.2 Example. Examples of isometries are translations, rotations, and reflections in a point or in a line or in a plane. ▬

V.2.3 Theorem. *1. An isometry on $\mathbb{R}^n$ mapping $0 \in \mathbb{R}^n$ to 0 is linear.*
2. The linear isometries on $\mathbb{R}^n$ are exactly the elements of $O_n(\mathbb{R})$.
3. Every isometry can be written as a composition of a translation and a linear isometry.
4. Isometries are invertible.

Proof. 1: One has $||u-v||^2 = \langle u-v,u-v\rangle = ||u||^2+||v||^2-2\langle u,v\rangle$ for $u,v \in \mathbb{R}^n$. If φ is an isometry and $\varphi(0) = 0$, then

$$\begin{aligned} 2\langle u,v\rangle &= ||u-0||^2+||v-0||^2-||u-v||^2 \\ &= ||\varphi(u)-\varphi(0)||^2+||\varphi(v)-\varphi(0)||^2-||\varphi(u)-\varphi(v)||^2 \\ &= 2\langle\varphi(u),\varphi(v)\rangle. \end{aligned}$$

A calculation now shows $||\varphi(u+av)-\varphi(u)-a\varphi(v)||^2 = 0$ for $a \in \mathbb{R}$, so φ is linear.
2: The proof of 1) shows that a linear isometry preserves the inner product, so is in $O_n(\mathbb{R})$. Vice versa, an element $A \in O_n(\mathbb{R})$ is clearly linear. It is an isometry since $||v-w||^2 = \langle v-w,v-w\rangle = \langle A(v-w),A(v-w)\rangle = ||A(v-w)||^2 = ||A(v)-A(w)||^2$. 3: Let φ be an isometry. Write $v = \varphi(0)$. Define $\tau_v : \mathbb{R}^n \to \mathbb{R}^n$ to be translation by v, so $\tau_v(w) = v+w$ for $w \in \mathbb{R}^n$. Moreover define $\psi : \mathbb{R}^n \to \mathbb{R}^n$ by $\psi(w) = \varphi(w)-v$. Then τ_v,ψ are isometries, and $\psi(0) = 0$ so from 1) we see ψ is linear. One has $\varphi(w) = \varphi(w)-v+v = \psi(w)+v = \tau_v(\psi(w))$ for all $w \in \mathbb{R}^n$, so $\varphi = \tau_v \circ \psi$.
4: By 3) and the fact that composing bijections yields a bijection, it suffices to show that translations and linear isometries are invertible. This is clear for translations, and the proof of Theorem V.1.2 shows it for linear isometries. ■

Because of this theorem linear isometries are the same as the orthogonal (linear) maps studied in Linear Algebra. For $\mathbb{R}^2$ we determined these maps in Example V.1.5: all reflections in a line through the origin, and the rotations with center the origin. We refer to the Appendix of these lecture notes for the case of $\mathbb{R}^3$ (the argument presented there, in fact works for $\mathbb{R}^n$ in general).

V.2.4 Definition. For $F \subset \mathbb{R}^n$, the set of all isometries on $\mathbb{R}^n$ with the property that F is mapped to F, forms a group. Namely, one easily verifies that the set is a subgroup of the group of *all* isometries. The resulting group is called *the symmetry group of* F.

It turns out that up to isomorphism the symmetry group of a set F is not affected by the *position* of F in $\mathbb{R}^n$, only by 'the shape of F':

V.2.5 Theorem. *If $F \subset \mathbb{R}^n$, $a \in \mathbb{R}_{>0}$, and φ is an isometry on $\mathbb{R}^n$, then the symmetry group of $a\varphi(F)$ and of F are isomorphic.*

Proof. The map $\sigma \mapsto a\varphi\sigma\varphi^{-1}\frac{1}{a}$ sends the symmetry group of F to that of $a\varphi(F)$ (compare Exercise 4), and this map is a homomorphism. it is bijective with inverse given by $\tau \mapsto \varphi^{-1}\frac{1}{a}\tau a\varphi$. ∎

We we now describe the symmetry groups of certain subsets of $\mathbb{R}^2$. In the Appendix of these lecture notes this is done for some subsets of $\mathbb{R}^3$ as well. A role will be played by the following result.

V.2.6 Lemma. *If G is a subgroup of $SO(2)$ consisting of exactly N elements, then G consists of all rotations over multiples of $2\pi/N$. In particular $G \cong \mathbb{Z}/N\mathbb{Z}$.*

Proof. Every element of SO(2), so in particular every element of G, is a rotation. Let $\sigma \in G$ be the rotation over over the smallest possible positive angle $2\pi\alpha$. Since G is finite, $\sigma^n = \mathrm{id}$ for some $n > 0$, so $n \cdot 2\pi\alpha$ is an integral multiple of 2π. This implies $\alpha \in \mathbb{Q}$, so we may write $\alpha = a/b$ for positive integers a, b with $\gcd(a,b) = 1$. Take $c, d \in \mathbb{Z}$ with $ac + bd = 1$, then σ^c is the rotation over $2\pi ac/b = 2\pi(1 - bd)/b$, i.e., over an angle $2\pi/b$. Since $2\pi a/b$ is the least positive angle of rotation in G, we have $a = 1$. We now show $b = N$. Take an arbitrary rotation $\tau' \in G$ over an angle $2\pi\ell/m$. By the same reasoning used above, we find a power τ of τ' representing rotation over $2\pi/m$. Note that τ' is a power of σ if and only if τ some power of σ. By taking a suitable combination $\sigma^p\tau^q$ we find an element of G representing rotation over $2\pi/\mathrm{kgv}(b,\ell)$. The minimality of $2\pi/b$ shows $\mathrm{lcm}(b,\ell) \leq b$, so $\ell | b$. This implies that every element of G is a power of σ. So $N = \#G = \mathrm{ord}(\sigma) = b$, proving the lemma. ∎

An alternative more geometric proof runs as follows. Take a circle around the origin and a point on it. The images of this point under the elements of G yield N points on the circle. Using that G consists of isometries, one can show that these N points are the vertices of a regular N-gon. The rotations permuting these vertices now form the group G.

V.3 The dihedral groups.

Let $C_r \subset \mathbb{R}^2$ be the circle with radius r around the origin. An isometry mapping C_r to itself necessarily fixes the center: namely, any point of C_r has a unique point of C_r at distance $2r$ (the antipodal point). As a result, symmetries of C_r will map lines through the origin to lines through the origin, and therefore the intersection point of these lines will be fixed. We conclude that the symmetry group of the circle is isomorphic to the group O(2).

V.3.1 Definition. The symmetry group of the circle C_r we call the *infinite dihedral group*. This group is denoted D_∞.

V.3.2 Theorem. *The group D_∞ is isomorphic to $O(2)$, and consists of reflections σ across arbitrary lines through the center of the circle, and of all rotations ρ around the center of the circle. The subset $R \subset D_\infty$ of all rotations is a commutative subgroup of D_∞.*

If $\sigma \in D_\infty$ is a reflection, then

$$D_\infty = R \cup R \cdot \sigma.$$

Taking σ the reflection across the x-axis, we have $\sigma\rho\sigma = \rho^{-1}$ for any $\rho \in R$.

Proof. We already saw that $D_\infty \cong \mathrm{O}(2)$ and that $\mathrm{O}(2)$ consists of reflections and rotations. The rotations are the matrices in $\mathrm{O}(2)$ of determinant 1, so R is the kernel of the homomorphism "determinant": $\mathrm{O}(2) \to \{\pm 1\}$.

The elements in $\mathrm{O}(2)$ having determinant -1 are reflections (their characteristic polynomial has the form $X^2 - t * X - 1$ for some $t \in \mathbb{R}$, so we have two real eigenvalues. They have absolute value 1 (since the matrix is orthogonal) and product -1. Hence the matrix represents the reflection across the line spanned by the eigenvector with eigenvalue $+1$).

The partition $D_\infty = R \cup R \cdot \sigma$ is the partition of D_∞ into rotations (determinant 1) and reflections (determinant -1).

The reflection σ in the x-axis is given by the matrix $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. For an arbitrary rotation $\rho = \left(\begin{smallmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{smallmatrix}\right)$ one computes

$$\sigma\rho\sigma = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} = \rho^{-1}$$

since $\cos(-\alpha) = \cos(\alpha)$ and $\sin(-\alpha) = -\sin(\alpha)$. This proves the theorem. ∎

When computing in D_∞ it is often convenient to regard rotations and reflections as maps on the complex plane $\mathbb{C}$. "Reflection in the x-axis" then becomes complex conjugation

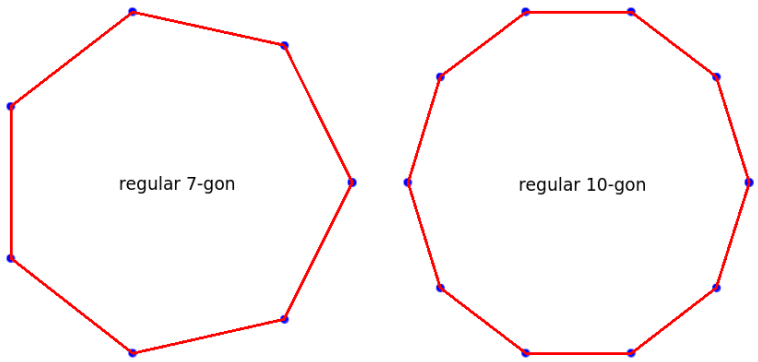
$$c : z \mapsto \overline{z}$$

and "rotation over α" becomes multiplication by $e^{\alpha i}$, so

$$m : z \mapsto e^{\alpha i} \cdot z.$$

As an example, cmc is the map sending z to $cmc(z) = cm(\overline{z}) = c(e^{\alpha i}\overline{z}) = \overline{e^{\alpha i}\overline{z}} = e^{-\alpha i}z$, so $cmc = v^{-1}$ as we saw earlier.

Subdividing the circle into $n \geq 2$ equal segments yields n vertices which define a regular n-gon F_n.



V.3.3 Definition. The symmetry group of F_n is called the n-th dihedral group D_n.

The center of F_n is fixed under all symmetries, so D_n is a subgroup of $D_\infty = \mathrm{O}(2)$. The group D_n consists of rotations and reflections; the rotations in D_n are those over an angle $k \cdot 2\pi/n$, for $0 \le k < n$. There are n of these. The rotation over the least positive angle $2\pi/n$ we denote by ρ. The reflections in D_n are precisely the reflections in either lines containing the origin and a vertex of F_n, or lines containing the origin and the midpoint of an edge of F_n. One of these reflections is de reflection in the x-axis, which from now on we denote by σ. There are precisely n reflections in D_n, namely all $\sigma\rho^k$ for $0 \le k < n$. So D_n is a finite group consisting of $n + n = 2n$ elements.

We summarize this discussion as follows.

V.3.4 Theorem. *The group D_n consists of $2n$ elements. For $n > 2$ the group D_n is non-commutative.*

Contained in D_n are the rotation ρ over an angle $2\pi/n$ and the reflection σ in the x-axis. Every element of D_n can be wriiten in a unique way as ρ^k or $\sigma\rho^k$, for some $0 \le k < n$.

One has $\mathrm{ord}(\rho) = n$ *and* $\mathrm{ord}(\sigma\rho^k) = 2$*, so in particular $\rho^n = \sigma^2 = id$. Moreover, $\sigma\rho\sigma = \rho^{-1}$.*

The subgroup R_n of D_n consisting of all rotations is isomorphic to $\mathbb{Z}/n\mathbb{Z}$.

Proof. The inverse ρ^{-1} is a rotation over an angle $(n-1)2\pi/n$. For $n > 2$ this differs from a rotation over $2\pi/n$, so then $\sigma\rho\sigma = \rho^{-1} \neq \rho$. This implies that D_n is not commutative if $n > 2$.

The remaining assertions in the theorem are evident, and/or are immediate consequences from Theorem V.3.2. ∎

V.3.5 Example. For $n = 2$ the group D_2 consists of 4 elements. In this case ρ is the map "rotate over 180 degrees", so $\rho(x, y) = (-x, -y)$. In particular $\rho^{-1} = \rho$, hence $\sigma\rho = \rho\sigma$. Therefore, D_2 is commutative. We knew this, because all groups consisting of only 4 elements are commutative. In the present case $\sigma\rho$ is the reflection across the y-axis. All elements $\neq id$ in D_2 have order 2 which implies $D_2 \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

Note that this example is somewhat odd: a regular 2-gon is just a line segment. The reflection across the line containing this segment is an element of order 2 in D_2. However, this reflection fixes every point in the segment F_2. ▬∎

V.3.6 Example. Let us determine the integers $n > 0$ for which the rotation r over 180 degrees given by $r(x, y) = (-x, -y)$, is an element of D_n.

We know that $\mathrm{ord}(r) = 2$. If $r \in D_n$ is a rotation, then $r = \rho^k$ for some k. Therefore, $r^n = \rho^{nk} = id$, so $2 = \mathrm{ord}(r)$ is a divisor of n and thus n is even. Vice versa, let $n = 2m$ for some integer $m > 0$. Then ρ^m is a rotation with $\mathrm{ord}(\rho^m) = 2$.Thus, ρ^m is a rotation over 180 degrees: $\rho^m = r$. Conclusion:

$$r \in D_n \Leftrightarrow n \text{ is even.}$$

Note that r is in the *center* of D_{2m}: $r\tau = \tau r$ for all $\tau \in D_{2m}$. ▬∎

V.4 Symmetries of a strip: frieze groups

This section discusses the symmetries of a strip in the plane, i.e., of the set of points in $\mathbb{R}^2$ located between two parallel lines. In particular so-called *discrete* subgroups

of this symmetry group will be presented. It turns out that this topic is related to art and to architecture.

Using Theorem V.2.5 the width of the strip can be scaled without changing the symmetry group in an essential way. Moreover we may change the strip by applying an isometry. This observation shows that the following choice of strip and group describes in some sense all cases.

V.4.1 Definition. By G_S we denote the group of symmetries of the set $S \subset \mathbb{R}^2$ defined by

$$S := \{(x,y) \in \mathbb{R}^2 \mid -1 \le y \le 1\}.$$

We start by presenting a more explicit description of the group G_S.

V.4.2 Theorem. The group G_S of all symmetries of the strip S consists of all isometries $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ given by $\varphi(x,y) = (\pm x + u, \pm y)$, for all four possibilities of the signs $(\pm,\pm)$ and all $u \in \mathbb{R}$.

Proof. Using Theorem V.2.3 and Example V.1.5 we find that the isometries φ of $\mathbb{R}^2$ are given by $\varphi(x,y) = (ax+by+u, cx+dy+v)$ with $a,b,c,d,u,v \in \mathbb{R}$ satisfying $a^2+c^2 = 1 = b^2+d^2$ and $ab+cd = 0$. If φ sends S to S, then in particular it sends the x-axis to itself. This means that the second coordinate of $\varphi(x,0) = (ax+u, cx+v)$ equals 0 for every $x \in \mathbb{R}$. From this one concludes $c = v = 0$. Hence $a^2 = 1 = b^2 + d^2$ and $ab = 0$, which implies $b = 0$ and $a^2 = 1 = d^2$. So indeed φ has the required form, and it is easy to verify that all isometries of this form send S to S, so they are in G_S. ∎

Here are the 4 types of elements in the symmetry group of the strip S:

- $\tau_u : (x,y) \mapsto (x+u, y)$ is the *translation* by $(u,0)$. We have $\mathrm{ord}(\tau_u) = \infty$ except when $u = 0$; obviously τ_0 is the identity map, which has order 1.
- $\rho_u : (x,y) \mapsto (-x+u, y)$ is the *reflection* across the vertical line given by $x = \frac{1}{2}u$. Since ρ_u^2 is the identity map and ρ_u is not, one finds $\mathrm{ord}(\rho_u) = 2$. Put $\rho := \rho_0$, then $\tau_u \rho = \rho_u = \rho \tau_{-u}$. In particular, all of these reflections can be expressed as a product of ρ and a translation.
- $\gamma_u : (x,y) \mapsto (x+u, -y)$ is called a *glide reflection*; in the case $u = 0$ it is the reflection $\gamma = \gamma_0$ across the x-axis. In general it is the composition of this reflection and a translation by $(u,0)$: we have $\gamma\tau_u = \gamma_u = \tau_u\gamma$. Since $\gamma_u^2 = \tau_{2u}$ one finds $\mathrm{ord}(\gamma_0) = 2$ and $\mathrm{ord}(\gamma_u) = \infty$ whenever $u \neq 0$.
- $\pi_u : (x,y) \mapsto (-x+u, -y)$ is the *point reflection* with center $(\frac{1}{2}u, 0)$. Clearly $\mathrm{ord}(\pi_u) = 2$. In terms of the point reflection $\pi = \pi_0$ in the origin, we have $\pi\tau_{-u} = \pi_u = \tau_u\pi$.

In particular, the description above shows that every element of G_S can be written as a product of a translation and an element in $\{\mathrm{id}, \rho, \gamma, \pi\}$. Note that $\{\mathrm{id}, \rho, \gamma, \pi\}$ is in fact a commutative subgroup of G_S: each of these four elements is its own inverse, and $\rho\gamma = \pi = \gamma\rho$ and $\rho\pi = \gamma = \pi\rho$ and $\gamma\pi = \rho = \pi\gamma$. So in fact this subgroup is isomorphic to $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. We summarize and extend this discussion in the following result.

V.4.3 Lemma. *The sets $T = \{\tau_u \mid u \in \mathbb{R}\}$ and $H = \{\mathrm{id}, \rho, \gamma, \pi\}$ are subgroups of G_S. Every element $g \in G_S$ can be expressed in a unique way as $g = th$ for some $t \in T$ and some $h \in H$. Similarly, every $g \in G_S$ can be expressed in a unique way as $g = h't'$ for some $t' \in T$ and some $h' \in H$. In other words, $G_S = TH = HT$. The map $\varphi : G_S \longrightarrow H$ defined by $g = th \mapsto \varphi(g) := h$ is a surjective group homomorphism with kernel T.*

Proof. We already observed that $H \subset G_S$ is a subgroup. The fact that also $T \subset G_S$ is a subgroup is easy to verify. The discussion preceding this lemma shows that every element of G_S can be written as th and as $h't'$ for some $t, t' \in T$ and some $h, h' \in H$. Since $g(0,0) = th(0,0) = t(0,0)$ it follows that t is the translation over $g(0,0)$, hence indeed g uniquely determines t and therefore also $h = t^{-1}g$. Similarly, if $g = h't'$ then $g^{-1} = t'^{-1}h'$ which determines t'^{-1} and hence t', as well as h'.

To see that φ is a homomorphism, first observe, as shown before, that for every $h \in H$ and every translation $t \in T$ there is another translation $\tilde{t} \in T$ (which is either t itself or its inverse) such that $ht = \tilde{t}h$. Hence, with $g = th$ and $g' = t'h'$ one finds $g'g = t'(h't)h = t'\tilde{t}h'h$. As a consequence, $\varphi(g'g) = h'h = \varphi(g')\varphi(g)$, which shows that φ is a homomorphism. It is evidently surjective, and has kernel T. ∎

Using the notation for elements of G_S as introduced above, the next equalities in G_S are not hard to verify.

(a) $\tau_u\tau_v = \tau_{u+v}$ and in particular $\tau_u^{-1} = \tau_{-u}$.
(b) $\rho_u\rho_v = \tau_{u-v}$ and $\rho_u^{-1} = \rho_u$ and $\rho_u\tau_v = \rho_{u-v} = \tau_{-v}\rho_u$.
(c) $\gamma_u\gamma_v = \tau_{u+v}$ and $\gamma_u^{-1} = \gamma_{-u}$.
(d) $\gamma_u\tau_v = \gamma_{u+v} = \tau_v\gamma_u$ and $\gamma_u\rho_v = \pi_{u+v} = \rho_v\gamma_{-u}$.
(e) $\pi_u\pi_v = \tau_{u-v}$ and $\pi_u^{-1} = \pi_u$.
(f) $\pi_u\tau_v = \pi_{u-v} = \tau_{-v}\pi_u$ and $\pi_u\rho_v = \gamma_{u-v} = \rho_{-v}\pi_{-u}$ and $\pi_u\gamma_v = \rho_{u-v} = \gamma_{-v}\pi_u$.

V.4.4 Definition. A *frieze group* is a subgroup $F \subset G_S$ with the property $F \cap T \cong \mathbb{Z}$.

Note that if $F \subset G_S$ is a frieze group, then by fixing an isomorphism $f : \mathbb{Z} \to F \cap T$ we have $t := f(1)$ is a translation in F, so $t = \tau_u$ for some $u \in \mathbb{R}$. Using that f is a homomorphism we have $f(n) = t^n = \tau_{nu}$. By assumption *every* translation in F is obtained in this way, in other words: $F \cap T = \{\tau_{nu} \mid n \in \mathbb{Z}\}$ and $|u| > 0$ is the minimal positive number such that the translation $\tau_{|u|}$ over $(|u|, 0)$ is in the frieze group F.

V.4.5 Example. $F := \{\tau_n \mid n \in \mathbb{Z}\}$ is a frieze group. By definition it consists of the translations over all points $(n, 0)$ for $n \in \mathbb{Z}$.

Also $F' := \{\gamma_1^n \mid n \in \mathbb{Z}\}$ is a frieze group. The glide reflections $\gamma_{2m+1} = \gamma_1^{2m+1}$ for $m \in \mathbb{Z}$ are in F', as well as all translations $\tau_{2m} = \gamma_1^{2m}$ for $m \in \mathbb{Z}$. In this case an isomorphism $\mathbb{Z} \cong F' \cap T$ is provided by $m \mapsto \gamma_1^{2m}$. ∎

We now present an alternative definition of "frieze groups". To achieve this, we first define a particular type of subgroup of the group $\mathrm{Isom}(\mathbb{R}^n)$ of all isometries of $\mathbb{R}^n$.

V.4.6 Definition. A subgroup $G \subset \mathrm{Isom}(\mathbb{R}^n)$ is called *discrete* if for every $v \in \mathbb{R}^n$ the ball $B_v := \{w \in \mathbb{R}^n \mid d(v, w) \leq 1\}$ has the following property:

$$\{g \in G \mid g(B_v) \cap B_v \neq \varnothing\}$$

is finite.

To understand this definition, observe that B_v is the n-dimensional ball with radius 1 and center $v \in \mathbb{R}^n$. Its image $g(B_v)$ under any isometry of $\mathbb{R}^n$ equals the ball $B_{g(v)}$. These balls have an empty intersection precisely when $||v - g(v)|| > 2$. So G is discrete, precisely when every $v \in \mathbb{R}^n$ has the property that only finitely many $g \in G$ send v to a point $g(v)$ at distance less than or equal to 2 from v.

V.4.7 Example. Clearly every finite subgroup of $\mathrm{Isom}(\mathbb{R}^n)$ is discrete. The infinite dihedral group D_∞ discussed in Definition V.3.1 and Theorem V.3.2 is *not* discrete: the group is infinite, and every element of it fixes the origin.

The translations $\tau_n \in G_S$ over a point $(n, 0)$ with $n \in \mathbb{Z}$ define an infinite discrete subgroup of $\mathrm{Isom}(\mathbb{R}^2)$: namely, for any $v \in \mathbb{R}^2$ and any $n \in \mathbb{Z}$ we have $||v - \tau_n(v)|| = |n|$, so only $\tau_0, \tau_{\pm 1}$, and $\tau_{\pm 2}$ send v to a point at distance ≤ 2 from v. ∎

V.4.8 Theorem. *A subgroup $F \subset G_S$ is a frieze group if and only if F is infinite and discrete.*

Proof. $\Rightarrow$: Assume $F \subset G_S$ is a frieze group. By definition $F \cap T \cong \mathbb{Z}$, so $F \cap T$ is infinite and therefore F is infinite, too. To verify that F is discrete we will use the homomorphism $\varphi : G_S \to H = \{\mathrm{id}, \rho, \gamma, \pi\}$. Write $n := \#\varphi(F)$ and $K := F \cap T$. Let $f_1 = \mathrm{id}, \ldots, f_n$ be elements of F such that $\varphi(F) = \{\varphi(f_1) = \mathrm{id}, \ldots, \varphi(f_n)\}$. Then $F = Kf_1 \cup \ldots \cup Kf_n$. As we observed earlier, F being a frieze group implies that K consists of all translations τ_{mc} for some fixed $c > 0$ with m ranging over the integers. Now let $v \in \mathbb{R}^2$ and take $f \in F$. Write $f = \tau_{mc} f_i$ for some $m \in \mathbb{Z}, i \in \{1, \ldots, n\}$. Then $f(B_v) = B_{f_i(v)+m(c,0)}$ so evidently $f(B_v) \cap B_v \neq \varnothing$ is only possible for finitely many values of m. Therefore F is discrete.
$\Leftarrow$: Now we assume that $F \subset G_S$ is an infinite discrete subgroup. The argument above shows that $F = Kf_1 \cup \ldots \cup Kf_n$ with $K = F \cap T$ and $f_1, \ldots, f_n \in F$. Since F is infinite, at least one (and therefore, all) of the sets Kf_i are infinite, so $\#K = \infty$. We have that F is discrete, and therefore its subgroup K is discrete as well. So we have a discrete group K consisting of translations over points $(c, 0)$, and what remains to be shown is that $K \cong \mathbb{Z}$. The definition of discreteness applied to K and to the ball $B_{(a,0)}$ of radius 1 and center $(a, 0)$ shows, that K contains only finitely many translations τ_c such that $|c - a| \leq 2$. In other words, every (closed) interval of length 4 in $\mathbb{R}$ contains only finitely many $c \in \mathbb{R}$ such that $\tau_c \in K$. As K is not empty, this implies that we can take the smallest possible $c > 0$ with $\tau_c \in K$. Claim: $K = \{\tau_{mc} | m \in \mathbb{Z}\}$ and $m \mapsto \tau_{mc}$ is an isomorphism $\mathbb{Z} \cong K$. Namely, by definition $\tau_c \in K$ hence since K is a group, for all $m \in \mathbb{Z}$ also $\tau_{mc} = \tau_c^m \in K$. This shows $K \supset \{\tau_{mc} | m \in \mathbb{Z}\}$. Vice versa if $\tau_d \in K$ for some $d \in \mathbb{R}$, then write $\frac{d}{c} = \ell + \epsilon$ with $\ell \in \mathbb{Z}$ and $0 \leq \epsilon < 1$. We have $d = \ell c + \epsilon c$ and therefore $\tau_{\epsilon c} = \tau_d \tau_{-\ell c} \in K$. By definition c is the smallest positive number with $\tau_c \in K$, so $0 \leq \epsilon c < c$ implies $\epsilon = 0$. This shows $d = \ell c$ and $\tau_d \in \{\tau_{mc} | m \in \mathbb{Z}\}$, completing the proof. ∎

We now present a description of all frieze groups.

V.4.9 Theorem. *All frieze groups are of exactly one of the following 7 types:*

F1. Groups $\{\tau_{mc} | m \in \mathbb{Z}\}$ (for fixed $c > 0$) consisting of translations;
F2. Groups $\{\gamma_c^m | m \in \mathbb{Z}\}$ (fixed $c > 0$) consisting of glide reflections and translations;
F3. Groups $\{\tau_{mc} | m \in \mathbb{Z}\} \cup \{\gamma_{mc} | m \in \mathbb{Z}\}$ (fixed $c > 0$) consisting of glide reflections and translations, including the reflection $\gamma = \gamma_0$;
F4. Groups $\{\tau_{mc} | m \in \mathbb{Z}\} \cup \{\rho_{u-mc} | m \in \mathbb{Z}\}$ (fixed $c > 0$ and fixed u) consisting of translations and reflections in vertical lines;
F5. Groups $\{\tau_{mc} | m \in \mathbb{Z}\} \cup \{\pi_{u-mc} | m \in \mathbb{Z}\}$ (fixed $c > 0$ and fixed u) consisting of translations and point reflections;
F6. Groups $\{\gamma_{c/2}^n | n \in \mathbb{Z}\} \cup \{\rho_u \gamma_{c/2}^n | n \in \mathbb{Z}\}$ (fixed $c > 0$ and fixed u) consisting of translations, glide reflections, reflections in vertical lines, and point reflections, not containing the glide reflection $\gamma = \gamma_0$.
F7. Groups $\{\tau_{mc} | m \in \mathbb{Z}\} \cup \{\rho_{u+mc} | m \in \mathbb{Z}\} \cup \{\gamma_{mc} | m \in \mathbb{Z}\} \cup \{\pi_{u+mc} | m \in \mathbb{Z}\}$ (fixed $c > 0$ and fixed u) consisting of translations, glide reflections, reflections in vertical lines, and point reflections, including the glide reflection $\gamma = \gamma_0$.

Proof. Let $F \subset G_S$ be a frieze group. We use the restriction to F of the homomorphism $\varphi : G_S \to H = \{\mathrm{id}, \gamma, \rho, \pi\}$ which we will (also) denote $\varphi : F \to H$. Its kernel equals $F \cap T$ and in the proof of Theorem V.4.8 we saw that this kernel consists of the translations τ_d^m for some fixed $d > 0$ and all $m \in \mathbb{Z}$. The image $\varphi(F) \subset H$ is by Theorem III.3.4 a subgroup of H. By discussing the possibilities for $\varphi(F)$ one by one, all possible F's will arise. Lagrange's theorem III.2.7 asserts that $\#\varphi(F) | \#H$, so $\#\varphi(F)$ equals 1, 2, or 4. In the latter case $\varphi(F) = H$, in the first case $\varphi(F) = \{\mathrm{id}\}$. In all other cases $\varphi(F)$ contains besides the identity id exactly one other element, of order 2. So we have the following possibilities.

1. $\varphi(F) = \{\mathrm{id}\}$. In this case F contains only translations. The proof of Theorem V.4.8 shows that $c > 0$ exists with $F = \{\tau_{mc} | m \in \mathbb{Z}\}$, which is a group isomorphic to $\mathbb{Z}$.
2. $\varphi(F) = \{\mathrm{id}, \gamma\}$; if this happens, then we are in one of the following cases.
   (a) Every glide reflection $\gamma_u = \tau_u \gamma \in F$ has infinite order, i.e., it satisfies $u \neq 0$. Put $K = F \cap T$ which, as above, can also be written as $K = \{\tau_{mc} | m \in \mathbb{Z}\}$ for some $c > 0$. Any $\gamma_u \in F$ yields $\gamma_u^2 = \tau_{2u} \in K$. Hence $u = mc/2$ for some $m \in \mathbb{Z}$. Now write $m = 2q + r$ with $q \in \mathbb{Z}$ and $r \in \{0, 1\}$. Then $\tau_{-qc} \in F$ and therefore also $\tau_{-qc}\gamma_{mc/2} = \gamma_{rc/2} \in F$. We conclude that $r = 1$ since otherwise F would contain γ_0, contrary to the assumption. As a consequence $\gamma_{c/2} \in F$ and moreover the glide reflections in F are precisely all $\gamma_{mc/2}$ with m an odd integer. Observe that $\gamma_{c/2}^n$ equals the translation τ_{mc} in case $n = 2m$ is even, and equals the glide reflection $\gamma_{mc/2}$ in case $n = 2m - 1$ is odd. This means $F = \{\gamma_{c/2}^n | n \in \mathbb{Z}\} = \{\tau_{mc} | m \in \mathbb{Z}\} \cup \{\gamma_{c/2}\tau_{mc} | m \in \mathbb{Z}\}$. The map $n \mapsto \gamma_{c/2}^n$ yields $\mathbb{Z} \cong F$.
   (b) F contains a glide reflection of finite order, i.e., $\gamma \in F$. As before, take $c > 0$ such that $F \cap T = \{\tau_{mc} | m \in \mathbb{Z}\}$. Is γ_u any glide reflection in F, then also $\gamma\gamma_u = \tau_u$ is in F which means $u = mc$ for some integer m. The converse holds as well: given $m \in \mathbb{Z}$ we have $\gamma_{mc} = \gamma\tau_{mc} \in F$. So
   $$F = \{\tau_{mc} | \, m \in \mathbb{Z}\} \cup \{\gamma_{mc} | \, m \in \mathbb{Z}\},$$
   and $(m, \overline{0}) \mapsto \tau_{mc}$ and $(m, \overline{1}) \mapsto \gamma_{mc}$ defines an isomorphism $\mathbb{Z} \times (\mathbb{Z}/2\mathbb{Z}) \cong F$.
3. $\varphi(F) = \{\mathrm{id}, \rho\}$. Again, there is a $c > 0$ such that $F \cap T = \{\tau_{mc} | m \in \mathbb{Z}\}$. Take any reflection $\rho_u \in F$, then $\rho_u\tau_c\rho_u = \tau_{-c}$; in particular it follows that F is not abelian. If also $\rho_v \in F$ then $\rho_u\rho_v = \tau_{u-v} \in F$ so $u - v = mc$ and $\rho_v = \rho_u\tau_{mc} = \rho_{u-mc}$. Clearly all products $\rho_u\tau_{mc}$ with $m \in \mathbb{Z}$ are in F, so
   $$F = \{\tau_{mc} | \, m \in \mathbb{Z}\} \cup \{\rho_{u-mc} | \, m \in \mathbb{Z}\}$$
   which indeed defines an infinite discrete group.
4. $\varphi(F) = \{\mathrm{id}, \pi\}$. This case is almost identical to the one above: let $c > 0$ such that $F \cap T = \{\tau_{mc} | m \in \mathbb{Z}\}$. Take any point reflection $\pi_u \in F$, then $\pi_u\tau_c\pi_u = \tau_{-c}$; in particular it follows that F is not abelian. If also $\pi_v \in F$ then $\pi_u\pi_v = \tau_{u-v} \in F$ so $u - v = mc$ and $\pi_v = \pi_u\tau_{mc} = \pi_{u-mc}$. Clearly all products $\pi_u\tau_{mc}$ with $m \in \mathbb{Z}$ are in F, so
   $$F = \{\tau_{mc} | \, m \in \mathbb{Z}\} \cup \{\pi_{u-mc} | \, m \in \mathbb{Z}\}$$
   which indeed defines an infinite discrete group.
5. $\varphi(F) = H = \{\mathrm{id}, \rho, \gamma, \pi\}$. In this case the elements of F that are either translations or reflections in a vertical line form a subgroup $F' \subset F$. In part 3. above it such groups F' were described: it is nonabelian and its elements are given in terms of two elements τ_c and ρ_u with $\rho_u\tau_c\rho_u = \tau_c^{-1}$. The remainder of the argument is in the spirit of Case 2. above:
   (a) Assume $\gamma = \gamma_0 \notin F$. Let $\gamma_v \in F$ be a glide reflection. Then $\gamma_v^2 = \tau_{2v} \in F$. Exactly as in 2(a) above, the translations and glide reflections in F are precisely the powers $\gamma_{c/2}^n$ of $\gamma_{c/2} \in F$. Moreover $\rho_u\gamma_{c/2}\rho_u = \gamma_{c/2}^{-1}$. We claim that all elements of F can be expressed in terms of ρ_u and $\gamma_{c/2}$. This was shown above for the translations, glide reflections, and the reflections in vertical lines. If $\pi_w \in F$ is a point reflection then $\pi_w = \rho_u\rho_u\pi_w = \rho_u\gamma_{u-w}$, where $\gamma_{u-w} = \rho_u^{-1}\pi_w \in F$ is a power of $\gamma_{c/2}$, which shows the claim in the remaining case. In fact, we find
   $$F = \{\gamma_{c/2}^n | \, n \in \mathbb{Z}\} \cup \{\rho_u\gamma_{c/2}^n | n \in \mathbb{Z}\}$$
   which indeed defines an infinite discrete subgroup of G_S.

(b) Assume $\gamma = \gamma_0 \in F$. Note that $\gamma g = g\gamma$ for every $g \in G_S$, hence this holds in particular for all $g \in F$. Take $F' \subset F$ the subgroup of F consisting of all translations and all reflections across vertical lines. We know from the third part of this proof that $F' = \{\tau_{mc} | \ m \in \mathbb{Z}\} \cup \{\rho_{u-mc} | \ m \in \mathbb{Z}\}$ for some $c > 0$. As before, one shows that F is the disjoint union $F' \cup \gamma F'$: clearly this union is in F and the sets $F', \gamma F'$ are disjoint. Moreover if $g \in F$, then it is in F' in case g is either a translation or a reflection in a vertical line. In the case g that is a point reflection or a glide reflection then $\gamma g \in F'$ and hence $g = \gamma\gamma g \in \gamma F'$. We conclude

$$F = \{\tau_{mc} | \ m \in \mathbb{Z}\} \cup \{\rho_{u+mc} | m \in \mathbb{Z}\} \cup \{\gamma_{mc} | \ m \in \mathbb{Z}\} \cup \{\pi_{u+mc} | \ m \in \mathbb{Z}\}$$

which indeed is an infinite discrete subgroup of G_S.

This completes the description of all frieze groups.

■

V.4.10 Remark. Observe that the group(s) encountered in the cases 3. and 4. and 5(a) above, are very reminiscent of the dihedral groups whose properties are listed in Theorem V.3.4. Namely, there is an element $\sigma \in \{\rho_u, \pi_v\}$ or order 2, and an element $r \in \{\tau_c, \rho_{c/2}\}$ of infinite order (in the case of the dihedral group D_n instead a rotation of order n is taken), and $\sigma r \sigma = r^{-1}$.

V.4.11 Remark. At first sight the groups appearing in the cases 5(a) and 5(b) of the preceding argument may look exactly the same. However, the groups are not even isomorphic. Indeed, starting from a group $F_4 = \{\tau_{mc} | \ m \in \mathbb{Z}\} \cup \{\rho_{u-mc} | \ m \in \mathbb{Z}\}$ (case 3 of the theorem) one obtains a group as in 5(a) as $F_6 = F_4 \cup \gamma_v F_4$ in which $v \notin \mathbb{Z} \cdot c$. A group as in 5(b) is given by $F_7 = F_4 \cup \gamma F_4$. The center (compare Exercise III.4(6)) $\mathcal{Z}(F_6)$ equals {id} whereas $\mathcal{Z}(F_7) = \{\text{id}, \gamma\}$, and isomorphic groups have isomorphic centers as well.

As abstract groups different types of frieze groups can be isomorphic, as we saw in the proof above and in the Remarks following the proof. Namely, the groups in F1 and F2 are all isomorphic to $\mathbb{Z}$, although in F1 a group contains only translations whereas in F2 also glide reflections occur. In a similar way the groups in F4, F5, and F6 are all isomorphic, although the different types contain different kinds of symmetries of the strip. In total we found 7 types of frieze groups, but only 4 different groups up to isomorphism. Only the groups in F1, F2, and F3 are abelian. Those in F3 contain an element of order 2, hence they are not isomorphic to a group in F1 or F2. All groups in F4, F5, F6, F7 are nonabelian. We already saw that three of these cases yield isomorphic groups, and also that the case F7 leads to groups not isomorphic to the one in the other cases.

We conclude by presenting examples of drawings on the strip S which have as symmetry group a given frieze group. Obviously we cannot extend the drawing unboundedly to the left or to the right, but it should be obvious how this is done.

Type F1:

$\cdots$b b b b b b b b b b b b b b b b b b b b b b b b b b b b b b b$\cdots$

Type F2:

$\cdots$b p b p b p b p b p b p b p b p b p b p b p b p b p b p b p b p b p$\cdots$

Type F3:

$\cdots$C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C$\cdots$

Type F4:

⋯b d b d b d b d b d b d b d b d b d b d b d b d b d b d b d b d b d⋯

Type F5:

⋯b q b q b q b q b q b q b q b q b q b q b q b q b q b q b q b q b q⋯

Type F6:

⋯b p q d b p q d b p q d b p q d b p q d b p q d b p q d b p q d b p q d⋯

Type F7:

⋯x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x⋯

Many examples of patterns with a frieze group as group of symmetries can be found on the internet, including their appearance in decorative art. The websites
`http://www.maa.org/sites/default/files/images/upload_library/4/vol1/architecture/Math/seven.html`
and
`https://en.wikipedia.org/wiki/Frieze_group`
provide a first impression.

V.5 Automorphisms of a graph

We will consider the simplest and easiest type of graph here. In particular we restrict ourselves to finite graphs with at most one edge between its vertices. Moreover, we do not prescribe a direction for the edges of our graphs. In more advanced graph theory the graphs we use, are called "finite simple undirected graphs"; our terminology will be shorter:

V.5.1 Definition. A graph Γ is a pair (V,E), with V a nonempty finite set (the 'vertices' of the graph), and E a (possibly empty) set consisting of subsets $\{a,b\} \subset V$ (the 'edges' of the graph).

V.5.2 Remark. A graph is usually presented as a picture: we draw its vertices as points, and we connect vertices a,b by a line segment (or by a loop in case $a = b$) precisely in the case $\{a,b\}$ is in the set of edges of the graph. In many examples such a picture is only possible if we allow some of the line segments to intersect. By emphasizing the actual vertices of the graph, one can make sure that no confusion with the intersection points of line segments arises.

To a graph one associates a finite group as follows.

V.5.3 Definition. An automorphism of a graph $\Gamma = (V,E)$ is a permutation σ on its set of vertices V, with the property that for all $\{a,b\} \in E$ also $\{\sigma(a),\sigma(b)\} \in E$.
The set consisting of all automorphisms of Γ is denoted $\mathrm{Aut}(\Gamma)$.

V.5.4 Theorem. *For a graph Γ with n vertices,* $\mathrm{Aut}(\Gamma)$ *is a subgroup of S_n.*

Proof. Enumerate the vertices of the graph Γ as $1,2,\ldots,n$. It is evident that any $\sigma \in \mathrm{Aut}(\Gamma)$ corresponds to a permutation in S_n, so $\mathrm{Aut}(\Gamma) \subset S_n$. We now show that this subset is a subgroup. The identity is contained in it. If $\sigma \in \mathrm{Aut}(\Gamma)$, then σ maps elements of E to elements of E as $\sigma(\{i,j\}) := \{\sigma(i),\sigma(j)\}$. This yields an injective map

from E to E, and since E is finite this map is surjective as well. This means that in case $\sigma(k) = i$ and $\sigma(\ell) = j$ and $\{i,j\} \in E$, then also $\{k,\ell\} \in E$. The definition of $\mathrm{Aut}(\Gamma)$ therefore shows that if $\sigma \in \mathrm{Aut}(\Gamma)$, then $\sigma^{-1} \in \mathrm{Aut}(\Gamma)$ as well. To prove that a product of elements in $\mathrm{Aut}(\Gamma)$ yields an element of $\mathrm{Aut}(\Gamma)$ is much simpler so we leave it for the reader. This completes the proof. ∎

V.5.5 Example. The *complete* graph Γ_n on n vertices is by definition the graph consisting of n vertices $1,2,\ldots,n$, and edges all $\{i,j\}$ with $1 \le i \le j \le n$. In this example the requirement that an automorphism sends edges to edges yields no restriction, hence $\mathrm{Aut}(\Gamma_n) = S_n$. —∎

V.5.6 Example. Enumerate the vertices of a regular n-gon as $1,2,\ldots,n$ (say, counter clockwise). Regard this n-gon as a graph F_n, so with vertices $1,2,\ldots n$ and edges $\{1,2\},\{2,3\},\ldots,\{n-1,n\},\{n,1\}$. Then $\mathrm{Aut}(F_n) \cong D_n$: namely, every symmetry of the regular n-gon can be considered as an element of $\mathrm{Aut}(F_n)$, so $D_n \subset \mathrm{Aut}(F_n)$. Vice versa, if $\tau \in \mathrm{Aut}(F_n)$ sends the vertex 1 to i, then 2 is mapped to one of the two neighbours of i, and this determines τ uniquely. As a consequence, at most $n \cdot 2 = 2n$ possible τ exist. We found this number of elements in the subset D_n, so indeed $\mathrm{Aut}(F_n) \cong D_n$. —∎

V.6 Exercises

1. Show that $D_2 \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ and $D_3 \cong S_3$.
2. Prove that D_n is commutative if and only if $n \leq 2$.
3. By considering the elements of D_n as permutations on the vertices of a regular n-gon one obtains a map from D_n to S_n. Verify that this map is an injective homomorphism. Which elements of D_n are mapped to even permutations?
4. Take $v \in \mathbb{R}^n$ and denote translation over v as τ_v. Let $a \in \mathbb{R}$ with $a \neq 0$.
   (a) Verify that $a\tau_v \frac{1}{a}$ is again a translation.
   (b) Show that if φ is any isometry on $\mathbb{R}^n$, then so is $a\varphi\frac{1}{a}$.
5. Let F_7 be a frieze group of type F7, and let $F_4 \subset F_7$ be its subgroup consisting of all translations and all reflections in vertical lines.
   (a) Show that every element $g \in F_7$ can be written in a unique way as $g = ab$ with $a \in \{\mathrm{id}, \gamma\}$ and $b \in F_4$.
   (b) Prove that $(\overline{n}, b) \mapsto \gamma^n b$ defines an isomorphism $(\mathbb{Z}/2\mathbb{Z}) \times F_4 \cong F_7$.
6. Find the type (according to Theorem V.4.9) of the symmetry group of each of the following infinite patterns:
   (a) $\cdots$ s s s s s s s s s s s s s s s s s s s s s s s $\cdots$
   (b) $\cdots$ v v v v v v v v v v v v v v v v v v v v v v v $\cdots$
   (c) $\cdots$ r r r r r r r r r r r r r r r r r r r r r r r $\cdots$
7. Verify that exactly 20 distinct graphs with exactly 3 vertices exist, and that none of them has A_3 as automorphism group.
8. Determine the number of automorphisms and the group $\mathrm{Aut}(H)$, with H the graph consisting of 6 vertices and 5 edges, drawn as the capital letter 'H'.
9. A cube can be considered as a graph by taking its 8 vertices as vertices and its 12 sides as edges. Determine the automorphism group of this graph.
10. The groups considered in this chapter all consist of bijections on a certain set, where the bijections are required to preserve some additional structure on the set. A natural additional example is to take as the set some group G, and to consider the bijections $\tau : G \to G$ which preserve the group structure:

$$\mathrm{Aut}(G) := \{\tau \in S_G \mid \tau(gh) = \tau(g)\tau(h)\}.$$

    (a) Determine $\mathrm{Aut}(\mathbb{Z}/4\mathbb{Z})$.
    (b) Show that $\mathrm{Aut}(G)$ is a group for any group G.
    (c) Show that $G \to \mathrm{Aut}(G)$ defined by $g \mapsto \gamma_g$, with γ_g defined by $\gamma_g(h) = ghg^{-1}$, indeed maps G to $\mathrm{Aut}(G)$, and moreover it is a group homomorphism from G to $\mathrm{Aut}(G)$.

VI CONJUGATION, INDEX, AND SYLOW THEORY

After two chapters in which we focused on *examples* of groups, we now continue the general theory started in Chapter III. In particular we develop more theory on subgroups of (mostly finite) groups. We also discuss multiplication on the left or on the right by a fixed element in more detail.

VI.1 conjugation

Consider an arbitrary group G and fix elements $a, b \in G$. The maps from G to G defined as 'multiplication on the left by a' (so $x \mapsto ax$) and 'multiplication on the right by b' ($x \mapsto xb$) are bijections. Their composition given by $x \mapsto axb$, is therefore bijective as well. In general this composed bijection is not a homomorphism. Namely, a homomorphism maps the unit element to the unit element. Our bijection maps $e \in G$ to $aeb = ab$. This equals the unit element e precisely when b is the inverse of a, so $b = a^{-1}$.

VI.1.1 Definition. If G is a group and $a \in G$, then the bijection $\gamma_a : G \to G$ given by $\gamma_a(x) = axa^{-1}$ is called the *conjugation* by a.

VI.1.2 Theorem. *Given is a group G and $a, b \in G$.*

1. *The conjugation γ_a by a is an isomorphism : $G \cong G$.*
2. *The conjugations γ_a, γ_b satisfy $\gamma_a\gamma_b = \gamma_{ab}$.*
3. *The inverse of γ_a is $\gamma_{a^{-1}}$.*
4. *If H is a subgroup of G, then so is $\gamma_a(H) = aHa^{-1}$, and $H \cong aHa^{-1}$.*

Proof. 1: For $x, y \in G$ we have $\gamma_a(xy) = axya^{-1} = axa^{-1}aya^{-1} = \gamma_a(x)\gamma_a(y)$. So γ_a is a homomorphism. We already observed that γ_a is bijective, so it is an isomorphism.
2: If $x \in G$, then $\gamma_a\gamma_b(x) = \gamma_a(bxb^{-1}) = abxb^{-1}a^{-1} = (ab)x(ab)^{-1} = \gamma_{ab}(x)$. In other words, $\gamma_a\gamma_b = \gamma_{ab}$.
3: From 2) follows $\gamma_a\gamma_{a^{-1}} = \gamma_e = \gamma_{a^{-1}}\gamma_a$. Every $x \in G$ satisfies $\gamma_e(x) = exe^{-1} = x$, so $\gamma_e = \mathrm{id}$. This shows the assertion.
4: Since γ_a is a homomorphism and H is a group, $\gamma_a(H)$ is a group as well. γ_a is injective, so this also holds for its restriction to H. This restriction by definition has image $\gamma_a(H)$, hence $H \cong \gamma_a(H)$. ∎

VI.1.3 Example. If G is a commutative group, then conjugation by an arbitrary element of G is the identity map. So conjugation can only be of interest for non-abelian groups.

In Linear Algebra conjugating matrices plays a major role when changing the basis of a vector space over a field. ▬

VI.1.4 Definition. Two elements x, y in a group G are called *conjugate* if a conjugation γ_a for some $a \in G$ exists with $\gamma_a(x) = y$.

The *conjugacy class* of $x \in G$ is by definition the subset of G given by

$$C_x = \{y \in G \mid \text{there exists } a \in G \text{ with } \gamma_a(x) = y\}.$$

VI.1.5 Example. In an abelian group G every $x \in G$ satisfies $C_x = \{x\}$.

In S_3 the cycles (1 2) and (1 2 3) are *not* conjugate. Namely, all $\tau \in S_3$ satisfy $\tau(1\,2)\tau^{-1} = (\tau(1)\,\tau(2))$ and $\tau(1\,2\,3)\tau^{-1} = (\tau(1)\,\tau(2)\,\tau(3))$. It follows that (1 2) is conjugate to every 2-cycle, and (1 2 3) to every 3-cycle, but the two given cycles are not conjugate. ▬

VI.1.6 Example. The theory about Jordan normal forms in Linear Algebra shows that two matrices $A, B \in \mathrm{GL}_n(\mathbb{C})$ are conjugate if and only if they have the same Jordan form. For example, $\left(\begin{smallmatrix}2&1\\0&2\end{smallmatrix}\right)$ en $\left(\begin{smallmatrix}2&0\\0&2\end{smallmatrix}\right)$ are *not* conjugate, but $\left(\begin{smallmatrix}2&1\\0&1\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}2&0\\0&1\end{smallmatrix}\right)$ are. ▬

VI.1.7 Theorem. *In any group G, 'being conjugate' is an equivalence relation. This means that:*

1. *Every $x \in G$ is conjugate to itself, so $x \in C_x$.*
2. *If x is conjugate to y, then also y is conjugate to x (so $x \in C_y$ implies $y \in C_x$).*
3. *If $x \in C_y$ and $y \in C_z$, then $x \in C_z$.*
4. *G is the disjoint union of subsets C_x. This means every element of G is in some C_x, and is an element in both C_x and C_y, then $C_x = C_y$.*

Proof. 1: We have $x = \gamma_e(x)$, so $x \in C_x$ for all $x \in G$.
2: $x, y \in G$ satisfy $x = \gamma_a(y)$ precisely when $y = \gamma_{a^{-1}}(x)$. This shows the assertion.
3: Given is that $a, b \in G$ exist with $\gamma_a(y) = x$ and $\gamma_b(z) = y$. It now follows that $\gamma_{ab}(z) = \gamma_a\gamma_b(z) = \gamma_a(y) = x$, so $x \in C_z$.
4: Every $a \in G$ is in some C_x since by 1) $a \in C_a$. If $a \in C_x$ and $a \in C_y$, then $c, d \in G$ exist with $a = \gamma_c(x)$ and $a = \gamma_d(y)$. Any $z \in C_x$ can therefore be written as $z = \gamma_f(x) = \gamma_{fc^{-1}d}(y)$, so $z \in C_y$. The same argument with x, y interchanged shows $C_y \subset C_x$. So $C_x = C_y$. ■

VI.1.8 Example. We write S_n as a disjoint union of conjugacy classes. Take $\sigma \in S_n$. Write σ as product of disjoint cycles: $\sigma = (a_1 \ldots a_{\ell_1})(a_{\ell_1+1} \ldots a_{\ell_2}) \ldots (a_{\ell_{s-1}+1} \ldots a_{\ell_s})$. A permutation τ sending each of the a_i to i (and the remaining integers in $\{1, \ldots, n\}$ bijectively to $\{\ell_s + 1, \ldots, n\}$) yields $\tau\sigma\tau^{-1} = (1\,2 \ldots \ell_1)(\ell_1 + 1 \ldots \ell_2) \ldots (\ell_{s-1} + 1 \ldots \ell_s)$. We conclude that all products of disjoint $\ell_1, \ell_2 - \ell_1, \ldots, \ell_s - \ell_{s-1}$-cycles are conjugate. The conjugacy class only depends on the set $\{\ell_1, \ell_2 - \ell_1, \ldots, \ell_s - \ell_{s-1}\}$ (the "*cycle type*").. In particular the number of pairwise different conjugacy classes equals the number of *partitions* of n; this is the number of presentations $n = \sum n_i$ with n_i positive integers, where the order of n_i's is not taken into account. The number of partitions is denoted $p(n)$. So $p(2) = 2$ since $2 = 2$ and $2 = 1 + 1$, and $p(4) = 5$ $(4, 3+1, 2+2, 2+1+1, 1+1+1+1)$. ▬

VI.1.9 Example. Determining the conjugacy classes in the alternating group A_n is considerably more involved than the case of S_n. We consider $n \leq 5$ here. A_n is commutative for $n \leq 3$. So in these cases $C_\sigma = \{\sigma\}$ for all $\sigma \in A_n$.

A_4 consists of (1), three products of two disjoint 2-cycles, and eight 3-cycles. Write $\{3, 4\} = \{a, b\}$, then $\tau = (2\,a\,b) \in A_4$ satisfies $\tau(1\,2)(3\,4)\tau^{-1} = (1\,a)(b\,2)$. So all products of two disjoint 2-cycles in A_4 are conjugate. Conjugating (1 2 3) by all 12

elements in A_4 one finds $C_{(1\,2\,3)} = \{(1\,2\,3),(1\,3\,4),(1\,4\,2),(2\,4\,3)\}$. The remaining four 3-cycles form a conjugacy class as well.

A_5 consists of 3-cycles, 5-cycles, products of 2 disjoint 2-cycles, and the identity. All 3-cycles form one conjugacy class. Namely, is $\sigma = (a_1\,a_2\,a_3)$, and is τ a permutation sending a_i to i, then $\tau\sigma\tau^{-1} = (1\,2\,3)$. Our concern here is of course, whether we may choose τ in A_5. Multiplying τ on the left by the 2-cycle $(4\,5)$ does not change the property above, and it shows we may indeed assume τ to be even. The products of two disjoint 2-cycles are all conjugate as well. Namely, such a product σ fixes a unique positive integer $i \le 5$, and any $\tau\sigma\tau^{-1}$ then fixes $\tau(i)$ (and no other positive integer ≤ 5). Using a suitable even τ then shows σ is conjugate to a product fixing 5. So we find precisely the products of two disjoint 2-cycles that were discussed when finding the conjugacy classes in A_4. As we saw there, they are all conjugate. What remains are the 5-cycles. A 5-cycle σ can be written as $\sigma = (1\,a\,b\,c\,d)$, so there are 24 of them. We have $\tau\sigma\tau^{-1} = \sigma$ precisely when $(\tau(1)\,\tau(a)\,\tau(b)\,\tau(c)\,\tau(d)) = (1\,a\,b\,c\,d)$, which means that τ must be a power of σ. The powers of σ form a subgroup H of A_n consisting of 5 elements. Write A_n as a disjoint union of subsets $H\pi$, with $\pi \in A_n$. Is $\tau \in H\pi$ then $\tau\sigma\tau^{-1} = \pi\sigma\pi^{-1}$. Moreover $\pi_1\sigma\pi_1^{-1} = \pi_2\sigma\pi_2^{-1}$ if and only if $\pi_2^{-1}\pi_1 \in H$, i.e., $\pi_1 \in H\pi_2$. So for a fixed 5-cycle σ there are as many pairwise distinct elements $\tau\sigma\tau^{-1}$ as there are distinct sets $H\pi$. Of those, there are $\#A_5/\#H = 12$. As a result the set of 5-cycles consists of two conjugacy classes, each containing 12 elements. In total we find 5 conjugacy classes, consisting of $1, 20, 15, 12$, and 12 elements, respectively. ∎

The example above shows for A_5 a result which is true for groups in general:

VI.1.10 Theorem. *If G is a group and $a \in G$, then $N(a) = \{x \in G \mid \gamma_x(a) = a\}$ is a subgroup of G. If G is finite, then*

$$\#G = \#N(a) \cdot \#C_a.$$

Proof. Using $\gamma_e = \mathrm{id}$ and $\gamma_{x^{-1}} = \gamma_x^{-1}$ and $\gamma_{xy} = \gamma_x\gamma_y$, we see that $N(a)$ is a subgroup of G. The proof of Theorem III.2.7 shows that G is a disjoint union of subsets $g_iN(a)$, for some $g_i \in G$. Suppose G is finite. Each of the subsets $g_iN(a)$ has $\#N(a)$ elements, so the proof is complete if we can show that the number of g_i's equals $\#C_a$, i.e., a bijection $\{g_1, \ldots, g_i, \ldots\} \to C_a$ exists. We claim that $g_i \mapsto x_i = \gamma_{g_i}(a) \in C_a$ is bijective. Namely, is $x \in C_a$ then $x = \gamma_g(a)$ for some $g \in G$. Then $g \in g_iN(a)$ for some i, hence $g = g_ih$ with $h \in N(a)$, and $x = \gamma_g(a) = \gamma_{g_i}\gamma_h(a) = \gamma_{g_i}(a) = x_i$. So the map is surjective. Is $x_i = x_j$, then $g_j^{-1}g_iag_i^{-1}g_j = a$ hence $g_j^{-1}g_i \in N(a)$. It follows that $g_iN(a) = g_jN(a)$, so $g_i = g_j$. So the map is injective as well, completing the proof. ∎

As an application we determine some conjugacy classes in A_n.

VI.1.11 Theorem. *Let $n \ge 5$.*

1. *In A_n all 3-cycles are conjugate.*
2. *In A_n all products of two disjoint 2-cycles are conjugate.*

Proof. Let $\sigma = (1\,2\,3) \in A_n$. By definition the subgroup $N(\sigma) \subset A_n$ consists of all even permutations τ satisfying $\tau\sigma\tau^{-1} = \sigma$, which means $(\tau(1)\,\tau(2)\,\tau(3)) = (1\,2\,3)$. As a consequence τ is a power of $(1\,2\,3)$ times an even permutation on $\{4, 5, \ldots, n\}$. So $\#N(\sigma) = 3\cdot(n-3)!/2$. Note that here the condition $n \ge 5$ is used. Theorem VI.1.10 implies $\#C_\sigma = (n!/2)/(3\cdot(n-3)!/2) = 2\binom{n}{3}$. This is equal to the number of 3-cycles in A_n, hence because C_σ consists of 3-cycles we conclude that all 3-cycles are conjugate.

The same idea can be adapted to the case of a product of two disjoint 2-cycles. We leave it as a useful exercise to the reader. ∎

VI.1.12 Remark. Here is an alternative proof. Take a 3-cycle $(a\ b\ c)$. Choose a permutation τ with $\tau(1)=a, \tau(2)=b$, and $\tau(3)=c$. Conjugation by τ and by $\tau\cdot(4\,5)$ both map $(1\,2\,3)$ to $(a\ b\ c)$. Since one of τ and $\tau\cdot(4\,5)$ is even, $(1\,2\,3)$ and $(a\ b\ c)$ are conjugate in A_n.

The case of products of two disjoint 2-cycles can be treated similarly; the reader is recommended to verify this!

VI.2 index

If G is a group and $H\subset G$ a subgroup, then for $g_1,g_2\in G$ the sets g_1H and Hg_2 can be mapped bijectively to H, namely using multiplication on the left by g_1^{-1}, respectively on the right by g_2^{-1}. Hence for fixed H all sets g_1H and Hg_2 are bijective. In particular, as was used earlier, in case H is finite they have the same number of elements. Another property we have used, is the fact that either $g_1H=g_2H$ (and this holds precisely when $g_1g_2^{-1}\in H$), or $g_1H\cap g_2H=\varnothing$. Recall for yourself how these assertions are proven!

VI.2.1 Definition. For H a subgroup of a group G the *index* of H in G is equal to the number of disjoint subsets of the form Hg in G. The index is denoted $[G:H]$. Is the index not finite then we write $[G:H]=\infty$.

VI.2.2 Remark. Since 'taking the inverse' $\iota: G\to G$ is a bijection and $\iota(Hg)=\iota(g)H$, one may also define the index in terms of subsets of the form gH.

VI.2.3 Theorem. *If G is a finite group, then $[G:H]$ is finite for all subgroups H. Moreover,*

$$\#G=[G:H]\cdot\#H.$$

Proof. This was already shown in the proof of Theorem III.2.7. ∎

VI.2.4 Example. It is certainly possible that a subgroup of an infinite group G has finite index. For example taking $G=\mathbb{Z}$, the subgroups are the groups $n\mathbb{Z}$. For $n\neq 0$ we have $\mathbb{Z}=n\mathbb{Z}\cup(1+n\mathbb{Z})\cup\ldots\cup((n-1)+n\mathbb{Z})$, so $[\mathbb{Z}:n\mathbb{Z}]=n$. Moreover, $[\mathbb{Z}:0\mathbb{Z}]=\infty$. ⌟∎

VI.3 Sylow theory

VI.3.1 Definition. (After P.L.M. Sylow, Norwegian mathematician, 1832–1918.) Let G be a finite group with $\#G=p^n\cdot m$ for a prime p, with $n\geq 1$ and $\gcd(p,m)=1$. A *Sylow p-group* in G is a subgroup $H\subset G$ with $\#H=p^n$.

VI.3.2 Example. Take a prime p and $n,m\geq 1$ with $\gcd(p,m)=1$. Then the group $G=\mathbb{Z}/p^n\mathbb{Z}\times\mathbb{Z}/m\mathbb{Z}$ consists of p^nm elements. There exists exactly one Sylow p-group in G, namely $H=\mathbb{Z}/p^n\mathbb{Z}\times\{\overline{0}\}$. Indeed, H is a Sylow p-group in G. Is also $H'\subset G$ such a group, then take $h=(a \bmod p^n, b \bmod m)\in H'$. Its order $\mathrm{ord}(h)$ divides $\#H'=p^n$. Moreover $\mathrm{ord}(b \bmod m)\,|\,\mathrm{ord}(h)\,|\,p^n$, and also $\mathrm{ord}(b \bmod m)$ divides $\#\mathbb{Z}/m\mathbb{Z}=m$. So $\mathrm{ord}(b \bmod m)\,|\,\gcd(p^n,m)=1$, implying $b \bmod m=\overline{0}$. This shows $H'\subset H$. Since $\#H'=p^n=\#H$ we conclude $H'=H$. ⌟∎

VI.3.3 Theorem. *Let G be a group with $\#G=p^nm$ for some prime p and integers $n,m>0$ with $\gcd(p,m)=1$.*

1. *G contains a Sylow p-group.*
2. *The number of pairwise distinct Sylow p-groups in G is $\equiv 1 \bmod p$.*
3. *If H and H' are Sylow p-groups in G then $H' = \gamma_a(H)$ for some $a \in G$.*
4. *The number of pairwise distinct Sylow p-groups in G is a divisor of m.*

Proof. Let N be the number of pairwise distinct Sylow p-groups in G. We must show $N \neq 0$; this will follow if we show the assertion in 2) saying that $N \equiv 1 \bmod p$. To this end, we study the collection of all subsets of G consisting of p^n elements.

If $H \subset G$ is a Sylow p-group and $g \in G$ then Hg is one of the sets under consideration. Theorem VI.2.3 shows that for given H there are $\#G/\#H = m$ such sets Hg. If $x \in G$, then $xHg = Hg$ precisely when $xH = H$; and this holds if and only if $x \in H$. So from $V = Hg$ we retrieve H as the set of all $x \in G$ with $xV = V$.

Now suppose $V \subset G$ is a subset with p^n elements, and moreover $\{x \in G \mid xV = V\}$ (which is a subgroup of G) also contains p^n elements. Is $v \in V$ and $x \in H$, then $xV = V$ hence $xv \in V$. So $Hv \subset V$, and because $\#Hv = p^n = \#V$ we have $V = Hv$. Conclusion: every set V with the given properties has the form Hg, with $g \in G$ and H a Sylow p-group. In total there are $N \cdot m$ such sets.

For any remaining $V \subset G$ with $\#V = p^n$, put $G_V = \{x \in G \mid xV = V\}$ (which is a subgroup of G). Just as in the proof of Theorem III.2.7, $V \subset G$ is a disjoint union of subsets $G_V \cdot v$, each having $\#G_V \neq p^n$ elements, and $\#G_V | \#V = p^n$. So $\#G_V = p^k$ for some $k < n$. Writing $\mathscr{P}$ for the collection of all these remaining V, we have

$$\binom{p^n m}{p^n} = Nm + \#\mathscr{P}.$$

We claim that $\#\mathscr{P} \equiv 0 \bmod p$. Indeed, let $V \in \mathscr{P}$. For every $x \in G$ is $xV \in \mathscr{P}$ as well, since $\#xV = \#V = p^n$, and $g \cdot (xV) = xV$ precisely when $x^{-1}gxV = V$. So $G_{(xV)} = \gamma_{x^{-1}}(G_V)$, which implies that $\#G_V = \#G_{(xV)}$. Now $x, y \in G$ satisfy

$$xV = yV \Leftrightarrow y^{-1}xV = V \Leftrightarrow y^{-1}x \in G_V \Leftrightarrow xG_V = yG_V.$$

As a consequence, for a given $V \in \mathscr{P}$ precisely $[G : G_V]$ pairwise different sets $xV \in \mathscr{P}$ exist. $V \in \mathscr{P}$ implies that $[G : G_V] = p^n m/p^k \equiv 0 \bmod p$. In this way $\mathscr{P}$ is partitioned into subcollections each containing a multiple of p sets V. This shows $\#\mathscr{P} \equiv 0 \bmod p$.

We conclude

$$]\binom{p^n m}{p^n} = Nm + \#\mathscr{P} \equiv Nm \bmod p.$$

From $\gcd(m, p) = 1$ we know $\overline{m} = m \bmod p$ is a unit in $\mathbb{Z}/p\mathbb{Z}$, and therefore

$$N \bmod p = \overline{m}^{-1} \cdot \binom{p^n m}{p^n} \bmod p.$$

This shows that $N \bmod p$ depends only on p, n, and m, and *not* on the actual group G! In particular, choosing $G = \mathbb{Z}/p^n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ one finds using Example VI.3.2 that $N \bmod p = 1 \bmod p$. This proves 1) and 2).

We now show 3). Let H, H' be Sylow p-groups in G. As we know, there are $[G : H] = m$ distinct subsets gH in G, and G is their union. We partition this collection of of sets gH into two classes $\mathscr{H}_1, \mathscr{H}_2$ as follows: $gH \in \mathscr{H}_1$ if $hgH = gH$ for all $h \in H'$, and $gH \in \mathscr{H}_2$ otherwise. By construction $m = \#\mathscr{H}_1 + \#\mathscr{H}_2$. We show, in fact in a way quite similar to what was done for the collection $\mathscr{P}$, that $\#\mathscr{H}_2 \equiv 0 \bmod p$. Let $gH \in \mathscr{H}_2$. Then $H'' = \{h \in H' \mid hgH = gH\}$ is a subgroup of H'. The definition of $\mathscr{H}_2$ implies $H'' \neq H'$, so $p | [H' : H'']$. Given $gH \in \mathscr{H}_2$ and $h \in H'$, consider hgH. This is clearly an element of $\mathscr{H}_2$. Moreover, for $h_1, h_2 \in H'$ one finds

$$h_1 gH = h_2 gH \Leftrightarrow h_2^{-1} h_1 gH = gH \Leftrightarrow h_2^{-1} h_1 \in H'' \Leftrightarrow h_1 H'' = h_2 H''.$$

So there are $[H' : H'']$ distinct sets hgH when h runs over H'. This partitions $\mathscr{H}_2$ into disjoint subsets, each having a multiple of p elements. So $\#\mathscr{H}_2 \equiv 0 \bmod p$. This implies

$$\#\mathscr{H}_1 \equiv m \bmod p \neq 0 \bmod p,$$

hence in particular $\mathscr{H}_1$ is nonempty. So gH exists with $hgH = gH$ for all $h \in H'$. In other words: $g^{-1}hg \in H$ for all $h \in H'$, which means $H' \subset \gamma_g(H)$. This shows 3).

Finally 4). Take H a Sylow p-group in G. There are N such groups, and we showed that each of them can be written as $\gamma_g(H)$ for some $g \in G$. Define

$$N(H) = \{g \in G \mid \gamma_g(H) = H\}.$$

This is a subgroup of G, and writing $G = \cup g_i N(H)$ one finds (check the details yourself!) that the Sylow p-groups in G are exactly the $\gamma_{g_i}(H)$, and these are pairwise distinct. So $N = [G : N(H)]$. Since $H \subset N(H)$ we have $\#N(H) = [N(H) : H] \cdot \#H$ and

$$N = [G : N(H)] = \#G/\#N(H) = \#G/([N(H) : H] \cdot \#H)|\#G/\#H = m.$$

This finishes the proof. ∎

VI.3.4 Corollary. *For p prime and $n, m > 0$ with $\gcd(p, m) = 1$ it holds that*

$$\binom{p^n m}{p^n} \equiv m \bmod p.$$

Proof. The proof of Theorem VI.3.3 showed $\binom{p^n m}{p^n} \equiv Nm \bmod p$ and $N \equiv 1 \bmod p$. This implies the corollary. ∎

VI.3.5 Example. S_4 has $24 = 3 \cdot 8$ elements. By Theorem VI.3.3 the number of Sylow 3-groups in S_4 divides 8, and has the form $3k + 1$. So their number is either 1 or 4. Each subset $\{(1), (a\ b\ c), (a\ c\ b)\}$ with $a \neq b, b \neq c, c \neq a$ is such a subgroup, and this yields 4 of them.

The number of Sylow 2-groups in S_4 is odd and it divides 3. So we have 1 or 3 of them. If H is such a group then $\#H = 8$, so every element in H has an order dividing 8. As a consequence, only 4-cycles, 2-cycles, products of two disjoint 2-cycles, and the identity can possibly belong to H. There can be at most two 2-cycles in H, and if there are then they are disjoint. Namely, otherwise a product $(a\ b)(b\ c) = (a\ b\ c)$ would be in H, which is impossible. It is also not possible that H contain only one 2-cycle. Namely, in that case every subgroup $\sigma H \sigma^{-1}$ also contains only one 2-cycle, and since all 2-cycles are conjugate this means there are at least as many Sylow 2-groups in S_4 as there are 2-cycles. Since there are six 2-cycles and at most three Sylow 2-groups, this is impossible. So H contains either no, or exactly two (disjoint) 2-cycles.

The number of 4-cycles in H is even, because a 4-cycle differs from its inverse and either both or none of the pair is in H. The square of a 4-cycle is a product of two disjoint 2-cycles, and a 4-cycle and its inverse have the same square and all other 4-cycles do not have this same square. As a consequence H contains exactly one 4-cycle and its inverse. Namely, at least one because otherwise H would contain at most 6 elements. Not more than one because otherwise conjugating one by powers of the other one checks that *all* 4-cycles are in H, and then all products of two disjoint 2-cycles as well. In that case $\#H \geq 1 + 6 + 3 = 10$.

We conclude that H consists of the identity, a 4-cycle and its inverse, all three products of two disjoint 2-cycles, and two 2-cycles. A small computation yields three such (conjugate) groups. One is

$$\{(1), (1\,2\,3\,4), (1\,4\,3\,2), (1\,2)(3\,4), (1\,3)(2\,4), (1\,4)(2\,3), (1\,3), (2\,4)\}.$$

Here is a much less elaborate way to construct this example: the symmetry group D_4 of the square has 8 elements. These elements permute the vertices of the square. So D_4 can be considered a subgroup of S_4. Explicitly: in the x,y-plane take the square with vertices $(\pm 1, \pm 1)$. The vertex in the i-th quadrant is denoted i. Rotating counter clockwise over 90 degrees yields the permutation (1 2 3 4) on the vertices Reflection in de x-axis corresponds to (1 4)(2 3). Reflection in the diagonal $x + y = 0$ yields (1 3), et cetera. ▬

The next two results illustrate the possible use of Theorem VI.3.3.

VI.3.6 Theorem. *Suppose $p \neq q$ are primes with $p \not\equiv 1 \bmod q$ and $q \not\equiv 1 \bmod p$, and G is a group with $\#G = pq$. Then $G \cong \mathbb{Z}/pq\mathbb{Z}$.*

Proof. By Theorem III.2.10 any $g \in G$ satisfies $\mathrm{ord}(g) \in \{1, p, q, pq\}$. The only element of order 1 is $e \in G$. If $\mathrm{ord}(g) = p$ then $\langle g \rangle$ is a Sylow p-group in G. The number of such groups divides q, so it is 1 or q. Moreover the number is $\equiv 1 \bmod p$. Since $q \not\equiv 1 \bmod p$ the number of Sylow p-groups in G equals 1. Every element of order p in G is in this Sylow p-group, so there are at most $p - 1$ such elements. (In fact there are precisely $p - 1$ of them, but we will not need this.)

The same reasoning shows that at most $q - 1$ elements in G have order q. Since $1 + p - 1 + q - 1 < pq = \#G$, the group G must contain elements of order pq. Is $g \in G$ any such element then $\langle g \rangle = G$, and $g \mapsto 1 \bmod pq$ yields an isomorphism $G \cong \mathbb{Z}/pq\mathbb{Z}$. ■

VI.3.7 Example. Applying Theorem VI.3.6 with $p = 3$ en $q = 5$, it follows that up to isomorphism only 1 group exists with 15 elements, namely $\mathbb{Z}/15\mathbb{Z}$. ▬

VI.3.8 Theorem. (Augustin-Louis Cauchy, French mathematician, 1789–1857)
Is G a finite group, and is p a prime dividing the number of elements of G, then $g \in G$ exists with $\mathrm{ord}(g) = p$.

Proof. Take a Sylow p-group $H \subset G$. This exists because of Theorem VI.3.3, and $\#H = p^k$ with $k \geq 1$. Take $x \in H$ such that $x \neq e$. Then $\mathrm{ord}(x) \neq 1$, and $\mathrm{ord}(x) | p^k$. Hence $\mathrm{ord}(x) = p^\ell$ with $1 \leq \ell \leq k$. Then $g = x^{p^{\ell-1}}$ has $\mathrm{ord}(g) = p$, as required. ■

VI.3.9 Remark. A proof of Theorem VI.3.8 which does not use Sylow theory is sketched in Exercise 10 below.

VI.3.10 Example. The requirement in Theorem VI.3.8 that p divides the number of elements in the group G, is necessary because of Theorem III.2.10. The requirement that p is prime is necessary as well. For example, the group D_4 has order 8, yet no element of order 8 exists in this group. More generally, if a group G consists of n elements, then an element of order n exists if and only if $G \cong \mathbb{Z}/n\mathbb{Z}$. In particular, this implies G is abelian. So in a non-abelian group with n elements, no element of order n exists.

As another example, S_4 consists of 24 elements. The positive divisors of 24 are $\{1, 2, 3, 4, 6, 8, 12, 24\}$. The divisors that occur as order of some element in S_4, are $\{1, 2, 3, 4\}$. ▬

VI.4 Exercises

1. Determine the number of elements of all conjugacy classes in S_6.
2. Find the conjugacy classes in A_6, and determine for each of them the number of elements.
3. Prove the second assertion in Theorem VI.1.11.
4. In the group D_n we have ρ = 'rotate counter clockwise over $2\pi/n$', and σ = 'reflect in the x-axis'.
   (a) Show that $\sigma\rho\sigma = \rho^{-1}$.
   (b) Show that every $\tau \in D_n$ can be written as $\rho^a\sigma^b$, with $0 \le a < n$ and $0 \le b \le 1$.
   (c) Take n odd. Find a conjugacy class in D_n consisting of n elements, another one with 1 element, and show there are $(n-1)/2$ remaining classes C_τ each containing 2 elements.
   (d) Now take n even. Show that D_n has two conjugacy classes with 1 element, $(n-2)/2$ conjugacy classes with 2 elements, and two with $n/2$ elements.
5. Suppose that G is a finite group with $\#G = n$, and G contains precisely 3 conjugacy classes.
   (a) Show that $n = 1 + a + b$, with $1 \le a \le b$ and $a|n$ and $b|n$.
   (b) Find all solutions to the equation in (a). (E.g., divide by n, and verify that $b \le 3$ holds.)
   (c) Use that any non-commutative group with 6 elements is isomorphic to S_3, and show that $G \cong \mathbb{Z}/3\mathbb{Z}$ or $G \cong S_3$. Check that these groups indeed have precisely 3 conjugacy classes.
6. Given a group G, a subgroup $H \subset G$ and any $g \in G$, show $[G : H] = [G : \gamma_g(H)]$.
7. Show that a finite *abelian* group G contains for every prime p with $p|\#G$ a unique Sylow p-group.
8. For every prime p, find the number of Sylow p-groups in S_5.
9. This exercise describes the Sylow p-groups in S_6.
   (a) Show that Sylow p-groups in S_6 do not exist for $p > 5$.
   (b) Show that the Sylow 2-groups are isomorphic to $D_4 \times \mathbb{Z}/2\mathbb{Z}$. Show there are $\binom{6}{2} \cdot 3 = 45$ such groups. (Consider a Sylow 2-group in S_4, and $(5\,6) \in S_6$.)
   (c) Show that the Sylow 3-groups are isomorphic to $\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$, and that there are $\binom{6}{3}/2 = 10$ of them. (Use disjoint 3-cycles.)
   (d) Show that there are 36 Sylow 5-groups, isomorphic to $\mathbb{Z}/5\mathbb{Z}$.
10. This exercise provides an alternative proof of Theorem VI.3.8. The argument is due to the British/Canadian mathematician John McKay.
    Let G be a finite group and let p be a prime with $p|n = \#G$. Consider

$$\mathscr{D} = \{(a_1, \ldots, a_p) \in G \times G \times \ldots \times G \mid a_1 a_2 \ldots a_p = e\}$$

    and its subsets $\mathscr{D}_1 = \{(a_1, \ldots, a_p) \in \mathscr{D} \mid a_1 = a_2 = \ldots = a_p\}$ and $\mathscr{D}_2 = \mathscr{D} \setminus \mathscr{D}_1$.
    (a) Show that $\#\mathscr{D} = n^{p-1} \equiv 0 \bmod p$.
    (b) Show that if $(a_1, a_2, \ldots, a_p) \in \mathscr{D}$ then also $(a_2, \ldots, a_p, a_1)$ and $(a_p, a_1, a_2, \ldots)$ are in $\mathscr{D}$ and more generally, so are all elements obtained by cyclically permuting the entries in $(a_1, a_2, \ldots, a_p)$.
    (c) Show that if $(a_1, a_2, \ldots, a_p) \in \mathscr{D}_2$, then the p elements of $\mathscr{D}_2$ obtained by cyclically permuting the entries, are pairwise distinct.
    (d) Prove that $\#\mathscr{D}_2 \equiv 0 \bmod p$.

(e) Prove that $\#\mathscr{D}_1 \geq p$, and conclude that the number of elements in G of order p is congruent to $p-1$ modulo p. In particular, such elements exist.

(f) For which of the steps in the proof is it crucial that p is prime?

11. Let G be a group with $\#G = 6$.

(a) Explain that $a, b \in G$ exist with $\mathrm{ord}(a) = 2$ and $\mathrm{ord}(b) = 3$.

(b) Show that if a, b as in (a) satisfy $\gamma_b(a) = a$, then $\mathrm{ord}(ab) = 6$ and $G \cong \mathbb{Z}/6\mathbb{Z}$.

(c) Show that if a, b as in (a) satisfy $\gamma_b(a) \neq a$, then C_a consists of 3 elements all having order 2, and $G \cong S_3$.

VII NORMAL SUBGROUPS AND FACTOR GROUPS

Reviewing Chapter II from a group theoretic perspective, we constructed starting from $\mathbb{Z}$ and its subgroup $N\mathbb{Z}$ a new group. Namely, the *elements* of this new group are the *residue classes* $a+N\mathbb{Z}$. Theorem II.1.5 and the succeeding definition and remark show that the group law on $\mathbb{Z}$ (addition) gives rise to a group law (addition as well) on these residue classes. This chapter discusses an extension of this construction for arbitrary groups G, using a subgroup $H \subset G$. So in particular we study how one may set up calculations with the classes gH for $g \in G$. As it turns out, only for certain subgroups called 'normal subgroups', one obtains a group structure on the collection of sets $\{gH \mid g \in G\}$ similar to what was done with $\mathbb{Z}$ and $n\mathbb{Z}$. As an advice, review the material from Section II.1 before studying the present chapter.

VII.1 Normal subgroups

Given a group G and a subgroup H, Theorem VI.1.2 says that for $a \in G$ the conjugate $\gamma_a(H) = aHa^{-1}$ is also a subgroup of G. Moreover H and $\gamma_a(H)$ are isomorphic. In general $H \neq \gamma_a(H)$. For example $H = \{(1), (1\,2)\}$ is a subgroup of S_3. With $a = (1\,3)$ we find $\gamma_a(H) = \{(1), (2\,3)\} \neq H$.

VII.1.1 Definition. A subgroup H of a group G is called *normal* if $H = aHa^{-1}$ for all $a \in G$.

VII.1.2 Example. In a commutative group G every subgroup H is normal. Namely, in this case $aha^{-1} = aa^{-1}h = h$ for all $a \in G$ and $h \in H$, so $aHa^{-1} = H$. ∎

VII.1.3 Example. In the dihedral group D_n the rotations form a subgroup, and this one is normal. Namely, regarded as linear maps on $\mathbb{R}^2$ the rotations in D_n are exactly the elements of D_n with determinant 1. If ρ is a rotation and $a \in D_n$, then $\det(a\rho a^{-1}) = \det(a)\det(\rho)\det(a^{-1}) = \det(a)\det(a)^{-1} = 1$, so $a\rho a^{-1}$ is also a rotation. ∎

VII.1.4 Example. We determine all normal subgroups in S_4. Take a normal $H \subset S_4$ and let $\sigma \in H$. Since $H = \tau H \tau^{-1}$ for every $\tau \in S_n$, H contains the conjugacy class C_σ of σ. Using that S_4 is a disjoint union of conjugacy classes (Theorem VI.1.7), also H is a disjoint union of conjugacy classes in S_4. The conjugacy classes in S_4 have $1, 6, 8$, and 3 elements, respectively. Obviously $(1) \in H$, so $\#H$ is a sum of some of the integers in $\{1, 3, 6, 8\}$ in which 1 appears as a term. Moreover $\#H | \#S_4 = 24$ by Theorem III.2.7. This leaves us with only a few possibilities:

1. $\#H = 1$, so $H = \{(1)\}$. Indeed this is a normal subgroup.
2. $\#H = 1+3 = 4$, so $H = \{(1), (12)(34), (13)(24), (14)(23)\}$. This is indeed a normal subgroup in S_4.
3. $\#H = 1+3+8 = 12$. In this case H consists of the identity, the products of two disjoint 2-cycles, and all 3-cycles. So $H = A_4$ which is a normal subgroup in S_4.
4. $\#H = 1+3+8+12 = 24$, so $H = S_4$.

We see that although S_4 has many subgroups, yet apart from S_4 itself and $\{(1)\}$ only two 'real' normal subgroups exist in S_4. ▬

VII.1.5 Example. As we saw, A_4 is normal in S_4. More generally A_n is normal in S_n. This follows from the fact that for permutations σ, τ one has $\epsilon(\sigma) = \epsilon(\tau\sigma\tau^{-1})$. (Alternatively, conjugating a product of disjoint cycles yields a product of disjoint cycles of the same type as before. In particular this does not affect the sign, hence $\tau A_n \tau^{-1} = A_n$.) ▬

VII.1.6 Example. Let G be a finite group and $\#G = p^n m$ with p prime, $n \geq 1$, and $\gcd(p, m) = 1$. Take a Sylow p-group $H \subset G$. All conjugate groups aHa^{-1} for $a \in G$ are Sylow p-groups as well, and Theorem VI.3.3 states that we obtain all Sylow p-groups in this way. We conclude that H is normal in G if and only if there is only one Sylow p-group in G. In many instances this condition may be verified using the two divisibility properties for the number of Sylow p-groups, as given in Theorem VI.3.3. ▬

The following useful lemma we in fact already used and derived in a number of earlier situations.

VII.1.7 Lemma. *If H is a subgroup of a group G and $a, b \in G$ then $aH = bH$ if and only if $b^{-1}a \in H$.*

Proof. In the proof of Theorem III.2.7 it was shown that two sets aH, bH are either equal or disjoint. Now $e \in H$ hence $a = ae \in aH$, so $aH = bH$ is equivalent to $a \in bH$. In other words: $aH = bH$ if and only if $a = bh$ for some $h \in H$, which in turn is equivalent to $b^{-1}a = h \in H$. ■

VII.1.8 Theorem. *Let G be a group and $H \subset G$ a subgroup. The next four statements are equivalent:*

1. *H is normal in G.*
2. *Every $a \in G$ satisfies $aH = Ha$.*
3. *For all $a \in G$ one has $aHa^{-1} \subset H$.*
4. *For all $a, b, c, d \in G$ with $aH = cH$ and $bH = dH$ one also has $abH = cdH$.*

Proof. 1) implies 2): If $a \in G$, then $aHa^{-1} = H$. Multiplying this equality on the right by a yields $aH = Ha$.
2) implies 3): Take $a \in G$ and $h \in H$. The assumption $aH = Ha$ implies $ah = h_1 a$ for some $h_1 \in H$, so $aha^{-1} = h_1 \in H$ which is what had to be shown.
3) implies 4): By Lemma VII.1.7 it suffices to show that if $c^{-1}a, d^{-1}b \in H$, then also $(cd)^{-1}(ab) = d^{-1}c^{-1}ab \in H$. Write $c^{-1}a = h_1 \in H$. Assuming 3) yields that $h_2 := d^{-1}h_1 d \in H$. Then also $d^{-1}c^{-1}ab = d^{-1}h_1 dd^{-1}b = h_2 d^{-1}b \in H$, finishing the argument.
4) implies 1): Let $h \in H$ and $a \in G$. One has $hH = eH$, hence using the assumption 4) also $ha^{-1}H = ea^{-1}H = a^{-1}H$. Lemma VII.1.7 therefore implies $aha^{-1} \in H$. This shows $aHa^{-1} \subset H$. Applying the argument to $a^{-1} \in G$ we also have $a^{-1}Ha \subset H$, and therefore $h = a(a^{-1}ha)a^{-1} \in aHa^{-1}$, so $H \subset aHa^{-1}$. This shows $H = aHa^{-1}$, so H is normal in G.
This shows that the four assertions are equivalent. ■

The fact that the rotations in D_n and the even permutations in S_n are normal subgroups, is a special case of the following.

VII.1.9 Theorem. *If G is a group and $H \subset G$ a subgroup with $[G:H] = 2$, then $H \subset G$ is normal.*

Proof. The condition $[G:H] = 2$ means that $a \in G$ exists such that G is the disjoint union of H and Ha. Hence $Ha = G \setminus H$. Since $a \notin H$, also H and aH are disjoint subsets of G. Again using $[G:H] = 2$ it follows from Remark VI.2.2 that $aH = G \setminus H$, so $aH = Ha$. Every subset $bH, Hb \subset G$ either equals H (in case $b \in H$) or equals aH (in case $b \notin H$). So $bH = Hb$ for all $b \in G$, and Theorem VII.1.8 shows that H is normal in G. ∎

VII.2 Factor groups

Theorem VII.1.8 shows among other things that in case a subgroup H of a group G is normal, then the rule $(aH)\cdot(bH) = abH$ is well defined. This means: if we write $aH = cH$ or $bH = dH$ for some elements $c, d \in G$, then the result $abH = cdH$ remains the same. (And vice versa, if H is *not* normal, then in general the result *will* depend on the element in G used for describing the set aH.)

VII.2.1 Example. Take $G = S_3$ and $H = \{(1), (1\,2)\} \subset S_3$. Then H is a subgroup of G, but H is not normal in G. Put $a = (1\,3)$ and $b = (1\,2\,3)$. Then

$$aH = \{(1\,3)(1),\ (1\,3)(1\,2)\} = \{(1\,3),\ (1\,2\,3)\}$$

and

$$bH = \{(1\,2\,3)(1),\ (1\,2\,3)(1\,2)\} = \{(1\,2\,3),\ (1\,3)\}.$$

So $aH = bH$ (as is also clear from $b^{-1}a = (1\,2) \in H$ using Lemma VII.1.7). However $a^2H = H$ and $b^2H = (1\,3\,2)H = \{(1\,3\,2),\ (2\,3)\} \neq a^2H$, which shows that in this case a multiplication on the sets $\{gH\}$ as above is not possible. ▬

VII.2.2 Definition. Given a group G and a normal subgroup $H \subset G$, the *factor group* G modulo H, which we denote as G/H, is the group having as elements the sets aH for $a \in G$. The unit element is $H = eH$, and the group law is defined by $(aH)\cdot(bH) = abH$.

VII.2.3 Remark. The given multiplication on G/H is well defined because of Theorem VII.1.8, and this uses the fact that H is normal in G. Since G is a group, it easily follows that G/H defines a group as well. For example the inverse $(aH)^{-1}$ of an element $aH \in G/H$ equals $a^{-1}H$. Namely, $(aH)\cdot a^{-1}H = eH$, which by definition is the unit element in G/H.

VII.2.4 Remark. The definition of the notion 'index' shows that the total number of pairwise distinct sets aH equals $[G:H]$. So $\#(G/H) = [G:H]$. Is G a finite group, then Theorem VI.2.3 shows $\#(G/H) = [G:H] = \#G/\#H$.

VII.2.5 Example. $H = N\mathbb{Z}$ is normal in $G = \mathbb{Z}$. The factor group is the group $\mathbb{Z}/N\mathbb{Z}$. This example shows in particular that a factor group of some infinite group may be finite. ▬

VII.2.6 Example. Let $n \geq 2$ and take $H = A_n$ as normal subgroup in $G = S_n$. Since A_n has index 2 in S_n, the factor group S_n/A_n consists of two elements. In particular, $S_n/A_n \cong \mathbb{Z}/2\mathbb{Z}$ because up to isomorphism there is only one group consisting of two elements. We conclude that a factor group of a non-abelian group may be commutative.

The two elements of S_n/A_n are by definition two subsets of S_n. One consists of all *even* permutations (A_n), and the other of all *odd* permutations (so $(1\,2)A_n$). The group law in S_n/A_n is described by the rule 'even times even is even' and 'odd times even is odd' and 'even times odd is odd' and 'odd times odd is even'. ▬

VII.2.7 Theorem. *If H is a normal subgroup of a group G, then the factor group G/H is commutative if and only if for all $a, b \in G$ the element $a^{-1}b^{-1}ab$ is in H.*

Proof. By definition G/H is abelian if and only if $(aH)\cdot(bH) = (bH)\cdot(aH)$ for all $a, b \in G$. The latter holds if and only if $abH = baH$ or all $a, b \in G$. By Lemma VII.1.7 this last condition is equivalent to $a^{-1}b^{-1}ab = (ba)^{-1}ab \in H$ for all $a, b \in G$. This proves the theorem. ■

VII.2.8 Example. Let $n \geq 3$. As we noted, S_n/A_n is a commutative group. Hence by Theorem VII.2.7 for all permutations σ, τ the product $\sigma^{-1}\tau^{-1}\sigma\tau$ is even. This of course is also clear from the fact that the sign ϵ is a homomorphism from S_n to a commutative group (± 1). Since $(a\ b)^{-1}(a\ c)^{-1}(a\ b)(a\ c) = (a\ b\ c)$ for pairwise distinct a, b, c and since every element of A_n can be written as a product of 3-cycles, it follows that if $H \subset S_n$ is normal and moreover S_n/H is abelian, then $A_n \subset H$, so $H = A_n$ or $H = S_n$. ▬

VII.2.9 Theorem. *Let H be normal in a group G. The assignment*

$$\pi : G \longrightarrow G/H \ : \ g \mapsto gH$$

defines a surjective homomorphism from G to G/H with $\ker(\pi) = H$.

Proof. For $a, b \in G$ one has $\pi(ab) = abH = (aH)\cdot(bH) = \pi(a)\pi(b)$. So π is a homomorphism. Any element in G/H has the form aH for some $a \in G$. Then $\pi(a) = aH$ showing that π is surjective. Finally, $a \in G$ satisfies $a \in \ker(\pi)$ precisely when $aH = eH$, so by Lemma VII.1.7 precisely when $a \in H$. Hence $\ker(\pi) = H$, which completes the proof. ■

VII.2.10 Remark. The homomorphism π given in Theorem VII.2.9 is usually called the *canonical* homomorphism to a factor group.

VII.2.11 Theorem. *A subgroup H of a group G is normal precisely when H is the kernel of some homomorphism from G to another group.*

Proof. If H is the kernel of a homomorphism, then it is a good and not too difficult exercise in using the given definitions to verify that indeed H is normal in G.

Vice versa, if H is normal then Theorem VII.2.9 shows H is the kernel of the canonical homomorphism from G to G/H. ■

VII.3 Simple groups

VII.3.1 Definition. A group G is called *simple* if $\{e\}$ and G are the only normal subgroups in G.

VII.3.2 Remark. If G is a simple group, G' any group, and $f : G \to G'$ a homomorphism, then either f is injective or f is the constant map sending every element of G to the unit element of G'. Namely, the kernel of f is normal in G, hence $\ker(f) = \{e\}$ (implying that f is injective) or $\ker(f) = G$ (meaning that everything is mapped to the unit element). This property of simple groups indicates that 'being simple' is a strong property.

VII.3.3 Example. We determine all nontrivial finite abelian simple groups G. Is G such a group and p is a prime dividing $\#G$, then by Theorem VI.3.8 $a \in G$ exists with $\mathrm{ord}(a) = p$. The subgroup $\langle a \rangle$ is normal in G and $\neq \{e\}$ (any subgroup of an abelian group is normal). So G being simple implies $G = \langle a \rangle \cong \mathbb{Z}/p\mathbb{Z}$. Indeed $\mathbb{Z}/p\mathbb{Z}$ is simple, because a subgroup has as number of elements a divisor of p and p is prime. We conclude that up to isomorphism the groups $\mathbb{Z}/p\mathbb{Z}$ are the only nontrivial simple finite abelian groups. ∎

VII.3.4 Remark. One of the main results in the modern theory of finite groups is a complete list of all simple finite groups. The list consists of some infinite 'families of simple groups' (such as the $\mathbb{Z}/p\mathbb{Z}$'s for p prime), and 26 more groups not appearing in any of the families. These additional ones are called the 'sporadic groups'. This list together with various properties of the groups is described in the book by J. Conway et al., *Atlas of finite simple groups*. Oxford: Clarendon Press, 1985. A digital version containing similar information can be found at `http://brauer.maths.qmul.ac.uk/Atlas/v3/`. The proof that the list is complete involved an enormous amount of collaboration to which over a hundred mathematicians contributed. In particular the American mathematician Daniel Gorenstein (1923–1992) deserves credit for this.

The largest sporadic group goes by the intriguing name 'the Monster'. This group consists of

$$2^{46} \cdot 3^{20} \cdot 5^9 \cdot 7^6 \cdot 11^2 \cdot 13^3 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71$$

elements.

Other finite simple groups include for example the groups $\mathrm{PSL}_n(\mathbb{Z}/p\mathbb{Z})$, for $n \geq 2$ and p prime and $(n,p) \neq (2,2)$ and $(n,p) \neq (2,3)$. These are the factor groups G/H, with G the group $\mathrm{SL}_n(\mathbb{Z}/p\mathbb{Z})$ of all $n \times n$ matrices with coefficients in $\mathbb{Z}/p\mathbb{Z}$ and determinant $\overline{1}$, and H the subgroup consisting of all matrices $\overline{a}I$, with $\overline{a}^n = \overline{1}$. Proving that these groups are indeed simple will not be done in the present course. What we *will* show however, is that the groups A_n for $n \geq 5$ are simple.

VII.3.5 Example. We show that A_5 is simple. Let $H \subset A_5$ be normal. If $\sigma \in H$, then also $\tau\sigma\tau^{-1} \in H$ for all $\tau \in A_5$. Hence H contains the conjugacy class C_σ of σ in A_5. It follows that H is a union of such conjugacy classes. These classes are pairwise disjoint and they contain $1, 12, 12, 15$, and 20 elements, respectively (compare Example VI.1.9). So

$$\#H = 1 + 12a + 15b + 20c$$

with $a \in \{0,1,2\}$ and $b, c \in \{0,1\}$. Moreover $\#H | \#A_5 = 60$. It is not hard to show that these conditions imply $\#H = 1$ or $\#H = 60$. Hence indeed A_5 is simple. ∎

VII.3.6 Theorem. *A_n is a simple group for every $n \geq 5$.*

Proof. The idea of the proof below is to show using $n \geq 5$ that a normal subgroup $H \neq \{(1)\}$ in A_n contains a 3-cycle. Then as a consequence H contains the conjugacy class in A_n of this 3-cycle. By Theorem VI.1.11 this conjugacy class contains *all* 3-cycles. And therefore Theorem IV.4.4 shows $H = A_n$, so A_n is simple.

Let $n \geq 5$ and take $H \neq \{(1)\}$ normal in A_5. Take $\sigma \neq (1)$ in H. Put $\sigma = \sigma_1\sigma_2\ldots\sigma_r$ with the σ_i disjoint ℓ_i-cycles and $\ell_1 \geq \ell_2 \geq \ldots \geq \ell_r \geq 2$.

If $\ell_1 \geq 4$ then let $\sigma_1 = (a_1\ a_2\ \ldots\ a_{\ell_1})$ and take $\tau = (a_1\ a_2\ a_3) \in A_n$. Since $H \subset A_n$ is normal, also $\sigma' = \tau\sigma\tau^{-1} \in H$. We take a closer look at σ'. The numbers a_1, a_2, a_3 occur in σ_1 and not in $\sigma_2, \ldots, \sigma_r$. Hence $\tau\sigma_i\tau^{-1} = \sigma_i$ for $i \geq 2$ and $\tau\sigma_1\tau^{-1} = (\tau(a_1)\,\tau(a_2)\,\ldots\,\tau(a_{\ell_1})) = (a_2\,a_3\,a_1\,a_4\,\ldots\,a_{\ell_1})$. Therefore

$$\sigma' = \tau\sigma\tau^{-1} = (\tau\sigma_1\tau^{-1})(\tau\sigma_2\tau^{-1})\ldots(\tau\sigma_r\tau^{-1}) = (a_2\,a_3\,a_1\,a_4\,\ldots\,a_{\ell_1})\sigma_2\ldots\sigma_r.$$

Now $\sigma^{-1}\sigma' = (a_{\ell_1}\,\ldots\,a_1)(a_2\,a_3\,a_1\,a_4\,\ldots\,a_{\ell_1}) = (a_1\,a_3\,a_{\ell_1}) \in H$. So in this case ($\ell_1 \geq 4$) the group H contains a 3-cycle, and we are done.

If $\ell_1 = \ell_2 = 3$ then write $\sigma_1 = (a_1\ a_2\ a_3)$ and $\sigma_2 = (b_1\ b_2\ b_3)$. Conjugation by $\tau = (a_1\ a_2\ b_1) \in A_n$ yields $\sigma' = (a_2\ b_1\ a_3)(a_1\ b_2\ b_3)\sigma_3\ldots\sigma_r \in H$. Hence also $\sigma^{-1}\sigma' = (a_1\ b_1\ a_2\ b_3\ a_3) \in H$. Using the argument presented for $\ell_1 \geq 4$ to this 5-cycle one concludes that H also contains a 3-cycle, finishing this case.

If $\ell_1 = 3$ and $\ell_i < 3$ for $i \neq 1$ then $\sigma^2 \in H$ is a 3-cycle and again we are done.

The final case is that σ is a product of disjoint 2-cycles. Since $\sigma \in H \subset A_n$, the number of 2 cycles here is even. Write $\sigma = (a\ b)(c\ d)\sigma_3\ldots\sigma_r$. Conjugation by $(a\ b\ c)$ yields $\sigma' = (b\ c)(a\ d)\sigma_3\ldots\sigma_r \in H$, hence $\sigma\sigma' = (a\ c)(b\ d) \in H$. So H contains the conjugacy class in A_n of $(a\ c)(b\ d)$, which by Theorem VI.1.11 means that *all* products of two disjoint 2-cycles are in H. In particular $(1\,2)(4\,5)\cdot(4\,5)(2\,3) = (1\,2\,3) \in H$. As before, this finishes the proof. ∎

VII.4 Exercises

1. Given are two groups G_1 and G_2 and a homomorphism $\varphi : G_1 \to G_2$. Show that $\ker(\varphi)$ is normal in G_1.
2. Show that if G is a group and $H \subset G$ a subgroup and $N \subset G$ normal, then $N \cap H$ is a normal subgroup in H.
3. Given $H \subset A_4$ consisting of (1) and all products of two disjoint 2-cycles. Show that H is normal in A_4. Give all elements of A_4/H, and construct a multiplication table for the group A_4/H.
4. Find a normal subgroup in $\mathbb{Z}/2\mathbb{Z} \times A_n$ containing exactly two elements. Prove that $\mathbb{Z}/2\mathbb{Z} \times A_n \ncong S_n$ for $n \neq 2$.
5. Prove that up to isomorphism only one group consisting of 1001 elements exists, as follows:
   (a) Such G contains normal subgroups N_7, N_{11}, and N_{13} with $7, 11$, and 13 elements, respectively;
   (b) Find an injective homomorphism $G \to G/N_7 \times G/N_{11}$;
   (c) Conclude from Theorem VI.3.6 that G is commutative
   (d) Prove that G contains an element of order 1001, and conclude $G \cong \mathbb{Z}/1001\mathbb{Z}$.
6. Find the subgroups of D_4 and for the normal ones N also D_4/N.
7. Given groups G_1, G_2 with unit elements e_1, e_2, show that $H = G_1 \times \{e_2\}$ is normal in $G_1 \times G_2$, and $(G_1 \times G_2)/H \cong G_2$.
8. Consider $G = \{\left(\begin{smallmatrix} 1 & a \\ 0 & b \end{smallmatrix}\right) \mid b \neq 0\} \subset \mathrm{GL}_2(\mathbb{R})$.
   (a) Show that G is a subgroup of $\mathrm{GL}_2(\mathbb{R})$, but not a normal one.
   (b) Show that $H_1 = \{\left(\begin{smallmatrix} 1 & 0 \\ 0 & b \end{smallmatrix}\right) \mid b \neq 0\}$ is a subgroup of G, not a normal one.
   (c) Show that $H_2 = \{\left(\begin{smallmatrix} 1 & a \\ 0 & 1 \end{smallmatrix}\right) \mid a \in \mathbb{R}\}$ is a normal subgroup in G.
   (d) Verify that $b \mapsto \left(\begin{smallmatrix} 1 & 0 \\ 0 & b \end{smallmatrix}\right) H_2$ defines an isomorphism between the multiplicative group $\mathbb{R} \setminus \{0\}$ and G/H_2.
9. Let G be a group and $N \subset G$ a normal subgroup. Suppose that H is any subgroup of G containing N.
   (a) Show that N is also normal in H.
   (b) Show that H/N is a subgroup of G/N.
   (c) Show that if $X \subset G/N$ is a subgroup, then $Y = \{a \in G \mid aN \in X\}$ is a subgroup of G containing N, and $X = Y/N$.
   (d) Prove that if $Y \subset G$ is a subgroup containing N, then Y/N is normal in G/N if and only if Y is normal in G.
10. Fix $k \geq 2$. In a group G consider the subset H consisting of all products $a_1^k a_2^k \dots a_n^k$, for $a_1, \dots, a_n \in G$. Show that H is a normal subgroup in G, and that every element of G/H has order dividing k.
11. In an *abelian* group G put $H = \{a \in G \mid \mathrm{ord}(a) < \infty\}$. Show that H is a normal subgroup in G. Prove that the unit element eH is the only element of G/H whose order is finite.
12. Find an example of a *non-abelian* group G such that $H = \{a \in G \mid \mathrm{ord}(a) < \infty\}$ is not a subgroup of G (e.g., think of real invertible 2×2 matrices).
13. Let $n \geq 5$ and take k *odd* with $3 \leq k \leq n$.
    (a) Given is a group G and a nonempty $X \subset G$ with the property that for all $x \in X$ and all $a \in G$, also $axa^{-1} \in X$. Prove that
    $$H = \{x_1^{\pm 1} \cdot \ldots \cdot x_r^{\pm 1} \mid x_i \in X\}$$
    is a normal subgroup in G.

(b) Show that A_n contains an element of order k.

(c) Show that A_5 contains no elements of order 4 or 6.

(d) Use a) to show that every element of A_n can be written as a product of elements of order k.

VIII HOMOMORPHISM– AND ISOMORPHISM THEOREMS

In this chapter a number of rules for dealing with factor groups G/H are treated. Here a central theme is how to describe a homomorphism from G/H to another group. The problem with this is that an element gH in G/H can be given in many different ways: it is possible that $gH = g'H$ while $g \neq g'$. An important example of this phenomenon already appeared in Lemma II.3.1. Anyone thoroughly understanding this lemma, will not find the more general situation described below in Criterion VIII.1.2 very challenging.

VIII.1 homomorphisms starting from a factor group

We begin with a property of any homomorphism starting from a factor group. Afterwards this property will be used to construct such homomorphisms.

VIII.1.1 Theorem. *Let G and G' be groups and $H \subset G$ a normal subgroup, and*

$$\varphi : G/H \longrightarrow G'$$

a homomorphism. given the canonical homomorphism $\pi : G \to G/H$ (so $\pi(g) = gH$), the composition $\psi = \varphi \circ \pi$ is a homomorphism $G \to G'$.

This homomorphism ψ satisfies $H \subset \ker(\psi)$.

Proof. Check for yourself that any composition of homomorphisms is again a homomorphism. So in particular ψ is a homomorphism $G \to G'$.

Also the second assertion in the theorem relies on a generality: $\psi = \varphi \circ \pi$ implies $\ker(\pi) \subset \ker(\psi)$, and we know $\ker(\pi) = H$. ∎

VIII.1.2 Criterion. *Let H be a normal subgroup of a group G, and consider an arbitrary group G'. Constructing a homomorphism $\varphi : G/H \to G'$ is done using the following recipe:*

1. *First find a homomorphism $\psi : G \to G'$ satisfying $H \subset \ker(\psi)$.*
2. *For ψ as in (1) it holds that $\psi(g_1) = \psi(g_2)$ for all $g_1, g_2 \in G$ such that $g_1 H = g_2 H$. In other words: the rule $\varphi(gH) = \psi(g)$ yields a well defined map from G/H to G'.*
3. *$\varphi : G/H \to G'$ as in (2) is a homomorphism, and $\psi = \varphi \circ \pi$ where π is the canonical homomorphism $G \to G/H$.*

Proof. We first show that any ψ as in (1) satisfies $\psi(g_1) = \psi(g_2)$ in case $g_1H = g_2H$. This follows from the fact that by Lemma VII.1.7 $g_1H = g_2H$ implies $g_2^{-1}g_1 \in H$. Now $H \subset \ker(\psi)$ shows that $g_2^{-1}g_1 \in \ker(\psi)$, so $\psi(g_2^{-1}g_1) = e'$, the unit element of G'. As a consequence $\psi(g_2)^{-1}\psi(g_1) = e'$ and thus $\psi(g_1) = \psi(g_2)$ which is what we wanted to show.

Next we show that the given φ indeed is a homomorphism. Let g_1H, g_2H be elements of G/H. Then $\varphi(g_1H \cdot g_2H) = \varphi(g_1g_2H)$ (by definition of the group law in G/H), and moreover $\varphi(g_1g_2H) = \psi(g_1g_2)$ (this is the definition of φ). Now ψ is a homomorphism, so $\psi(g_1g_2) = \psi(g_1)\psi(g_2)$ which by the definition of φ equals $\varphi(g_1H)\varphi(g_2H)$. We conclude $\varphi(g_1H \cdot g_2H) = \varphi(g_1H)\varphi(g_2H)$, so φ is a homomorphism.

Finally, for arbitrary $g \in G$ we have $(\varphi \circ \pi)(g) = \varphi(gH) = \psi(g)$ so indeed $\psi = \varphi \circ \pi$. ∎

VIII.1.3 Example. We will determine all homomorphisms from $\mathbb{Z}/12\mathbb{Z}$ to $\mathbb{Z}/4\mathbb{Z}$. Using Theorem VIII.1.1 and Criterion VIII.1.2 this boils down to finding all homomorphisms from $\mathbb{Z}$ to $\mathbb{Z}/4\mathbb{Z}$ having $12\mathbb{Z}$ in the kernel. This condition does not provide any restriction. Namely, is $f : \mathbb{Z} \to \mathbb{Z}/4\mathbb{Z}$ an arbitrary homomorphism and $n \in 12\mathbb{Z}$, then $n = 12m$ with $m \in \mathbb{Z}$, so in particular $n = 3m + 3m + 3m + 3m$ implying $f(n) = f(3m) + f(3m) + f(3m) + f(3m) = \overline{0}$.

So we simply look for all homomorphisms from $\mathbb{Z}$ to $\mathbb{Z}/4\mathbb{Z}$. Any such homomorphism sends the unit element to the unit element, so 0 to $\overline{0}$. Suppose $\overline{a} \in \{\overline{0}, \overline{1}, \overline{2}, \overline{3}\}$ is the image of 1. This determines the map, because $1 + \ldots + 1$ must be mapped to $\overline{a} + \ldots + \overline{a}$, and the opposite of $1 + \ldots + 1$ to the opposite of $\overline{a} + \ldots + \overline{a}$. Verify yourself that in this way indeed a homomorphism is obtained.

In total we therefore find 4 pairwise different homomorphisms from $\mathbb{Z}/12\mathbb{Z}$ to $\mathbb{Z}/4\mathbb{Z}$. Each of them is completely determined by the image of 1 mod 12. ∎

VIII.1.4 Example. The group D_4 consisting of all symmetries of the square consists, as we have seen, of 8 elements. We place the square in the plane in such a way that its center is the origin. In this case the 8 symmetries given by linear maps $\mathbb{R}^2 \to \mathbb{R}^2$. The symmetry 'reflect in the origin' forms together with the identity map a subgroup $H = \{\pm 1\} \subset D_4$. It is not hard to verify that H is a normal subgroup of D_4. We now present an isomorphism $D_4/H \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

Criterion VIII.1.2 tells us to start constructing a homomorphism from D_4 to $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. The determinant provides a homomorphism $D_4 \to \pm 1 \cong \mathbb{Z}/2\mathbb{Z}$. Furthermore the square has 2 diagonals, and every element of D_4 permutes these two. This defines a second homomorphism $f : D_4 \to S_2 \cong \mathbb{Z}/2\mathbb{Z}$. The pair $\psi = (\det, f)$ is the requested homomorphism

$$\psi : D_4 \longrightarrow \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} : \sigma \mapsto (\det(\sigma), f(\sigma)) \in \pm 1 \times S_2 \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}.$$

Note that ψ is surjective (find yourself explicit elements in D_4 that are mapped to each of the elements of $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$). The kernel of ψ consists by definition of all symmetries fixing the two diagonals, and moreover having determinant 1. So the kernel consists of the rotations ± 1, in other words it is precisely our subgroup H. So the condition $H \subset \ker(\psi)$ in Criterion VIII.1.2 is satisfied. One concludes that a homomorphism $\varphi : D_4/H \to \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ exists, given by $\varphi(\sigma H) = \psi(\sigma)$. Since $\psi(\sigma)$ ranges over all elements of $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, it follows that φ is surjective. The number of elements in D_4/H equals $[D_4 : H] = \#D_4/\#H = 8/2 = 4$ which is also the number of elements of $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. Since φ is surjective this implies φ is a bijection and therefore it is an isomorphism. So $D_4/H \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, as asserted. ∎

The most frequently used rule for dealing with factor groups reads as follows.

VIII.2.1 Theorem. *If $\psi : G \to G'$ is a homomorphism of groups and $H = \ker(\psi)$, then H is a normal subgroup of G and*

$$G/H \cong \psi(G) \subset G'.$$

In particular for ψ surjective one has $G/H \cong G'$.

Proof. The fact that H is a normal subgroup is already stated in Theorem VII.2.11. Moreover by Criterion VIII.1.2 $\varphi(gH) = \psi(g)$ in our case results in a well defined homomorphism φ from G/H to G'.

We determine the kernel of φ. Note that $gH \in \ker(\varphi)$ precisely when $\varphi(gH)$ is the unit element $e' \in G'$. Now $\varphi(gH) = \psi(g)$ equals e' exactly when $g \in \ker(\psi) = H$. Moreover $g \in H$ is equivalent to $gH = eH$, i.e., gH is the unit element in G/H. The conclusion is that $\ker(\varphi)$ consists of only the unit element in G/H. Theorem III.3.6 therefore implies that φ is injective.

Injectivity of φ yields that G/H is isomorphic to the image of φ, and by definition this equals the image of ψ. So $G/\ker(\psi) \cong \psi(G)$, as required. In case ψ is surjective we have $\psi(G) = G'$ hence $G/\ker(\psi) \cong G'$. ∎

VIII.2.2 Example. The determinant is a surjective homomorphism from $\mathrm{GL}_n(\mathbb{R})$ to the multiplicative group $(\mathbb{R} \setminus \{0\}, \cdot, 1)$. The kernel of the determinant is $\mathrm{SL}_n(\mathbb{R})$, so Theorem VIII.2.1 implies that

$$\mathrm{GL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{R}) \cong \mathbb{R} \setminus \{0\}.$$

—∎

VIII.2.3 Example. The complex numbers $a + bi$ satisfying $a^2 + b^2 = 1$ are a subgroup $\mathbf{T}$ of the multiplicative group $(\mathbb{C} \setminus \{0\}, \cdot, 1)$. This subgroup is isomorphic to the factor group $\mathbb{R}/\mathbb{Z}$. Namely, $x \mapsto e^{2\pi i x}$ defines a surjective homomorphism $\mathbb{R} \to \mathbf{T}$ with kernel $\mathbb{Z}$. —∎

The next two isomorphism theorems for factor groups are in fact consequences of Theorem VIII.2.1.

VIII.2.4 Theorem. *Given are a group G, an arbitrary subgroup $H \subset G$, and a normal subgroup $N \subset G$. Then*

1. *$HN = \{hn \mid h \in H \text{ and } n \in N\}$ is a subgroup of G.*
2. *N is a normal subgroup of HN.*
3. *$H \cap N$ is a normal subgroup of H.*
4. *$H/(H \cap N) \cong HN/N$.*

Proof. 1: By Theorem III.2.3 we have to check the three conditions (H1, H2, H3). H1: from $e = e \cdot e$ and $e \in H, e \in N$ we see $e \in HN$. H3: for arbitrary $h \in H$ and $n \in N$ we know $hn^{-1}h^{-1} \in N$ since N is normal in G. So $(hn)^{-1} = n^{-1}h^{-1} = h^{-1} \cdot (hn^{-1}h^{-1})$ is in HN. Finally H2: for $h_1, h_2 \in H$ and $n_1, n_2 \in N$ we see $h_2^{-1}n_1h_2 \in N$ since $N \subset G$ is normal. Therefore $(h_1n_1) \cdot (h_2n_2) = h_1h_2(h_2^{-1}n_1h_2)n_2 \in HN$. So indeed $HN \subset G$ is a subgroup.
2: From $e \in H$ it follows that $N = eN \subset HN$. Since N is a group it is therefore a subgroup of HN. We have $gNg^{-1} = N$ for all $g \in G$, so certainly for those $g \in G$

which are in HN. Hence N is normal in HN.
3 and 4: define $\psi : H \to G/N$ by $\psi(h) = hN \in G/N$. This is the restriction to H of the canonical homomorphism $G \to G/N$, so ψ is a homomorphism. Note $h \in \ker(\psi)$ precisely when $hN = N$, which means when $h \in N$. So $\ker(\psi) = H \cap N$ which implies by Theorem VII.2.11 that $H \cap N$ is normal in H. Theorem VIII.2.1 tells us that $H/(H \cap N)$ is isomorphic to the image of ψ. So our argument is complete if we have shown $\psi(H) = HN/N$. This is straightforward: an element of $\psi(H)$ can be written as $hN \in G/N$, and here $h \in H \subset HN$ so this element is in HN/N. Vice versa, an element in HN/N can be written as hnN with $h \in H$ and $n \in N$. Now $nN = N$ hence $hnN = hN$, which is the image of $h \in H$ under ψ. This completes the proof. ∎

VIII.2.5 Example. Take $G = \mathbb{Z}$, $n, h \in \mathbb{Z}$ and $H = h\mathbb{Z}$ and $N = n\mathbb{Z}$. The group law in $\mathbb{Z}$ is 'addition'; therefore Theorem VIII.2.4 in the present case says that $h\mathbb{Z}/(h\mathbb{Z} \cap n\mathbb{Z}) \cong (h\mathbb{Z} + n\mathbb{Z})/n\mathbb{Z}$. We analyse this a bit further. Note that $h\mathbb{Z} \cap n\mathbb{Z}$ consists of all integers which are both divisible by h and by n. Corollary I.2.9 (5) asserts that these are precisely the multiples of $\mathrm{lcm}(h,n)$. So $h\mathbb{Z} \cap n\mathbb{Z} = \mathrm{lcm}(h,n)\mathbb{Z}$. Moreover Theorem I.1.11 implies $h\mathbb{Z} + n\mathbb{Z} = \gcd(h,n)\mathbb{Z}$. We conclude

$$h\mathbb{Z}/\mathrm{kgv}(h,n)\mathbb{Z} \cong \gcd(h,n)\mathbb{Z}/n\mathbb{Z}.$$

In the special case $\gcd(h,n) = 1$ this says $h\mathbb{Z}/hn\mathbb{Z} \cong \mathbb{Z}/n\mathbb{Z}$. Incidentally, this last isomorphy also holds for $\gcd(h,n) \neq 1$, as can (for example) be shown using Theorem VIII.2.1. ▬∎

VIII.2.6 Example. Consider $N = \{(1), (1\,2)(3\,4), (1\,3)(2\,4), (1\,4)(2\,3)\}$ which is a subgroup of $G = S_4$, and $H =$ all permutations in S_4 fixing the integer 4 (so $H = S_3 \subset S_4$). Note that N is normal in G. We have $HN = S_4$, since both H and N are subgroups of HN, so $\#HN$ is divisible by both $\#H = 4$ and $\#N = 6$. Hence $12 | \#HN$. As a consequence $[S_4 : HN] = 1$ or $= 2$. In particular HN is normal in S_4. Now S_4/HN has at most two elements, so this factor group is abelian, implying $A_4 \subset HN$. Noting that HN contains odd permutations as well, it follows that $HN = S_4$. (This may be verified in numerous other ways, but the argument above illustrates a number of techniques we now have at our disposal.) Observing that $H \cap N = \{(1)\}$, Theorem VIII.2.4 implies

$$S_4/H = HN/H \cong H/(H \cap N) = S_3/\{(1)\} = S_3.$$

▬∎

Part of the next result we encountered in Exercise 9 of Chapter VII.

VIII.2.7 Theorem. *Given is a group G and a normal subgroup $N \subset G$.*

1. *Every normal subgroup in G/N has the form H/N, with H a normal subgroup in G containing N.*
2. *Is $N \subset H$ for some normal subgroup H in G, then $(G/N)/(H/N) \cong G/H$.*

Proof. For (1) we refer to Exercise 9 in Chapter VII.
(2): Consider the canonical homomorphism $\pi : G \to G/H$. We have $N \subset H$ and by Theorem VII.2.9 $H = \ker(\pi)$, so $N \subset \ker(\pi)$. Hence applying Criterion VIII.1.2 one concludes that $\psi(gN) = \pi(g) = gH$ defines a homomorphism $\psi : G/N \to G/H$. This homomorphism ψ is surjective because π is surjective. Moreover $gN \in \ker(\psi)$ precisely when $\psi(gN) = gH = eH$, so $\ker(\psi)$ consists of all classes gN with $g \in H$. We conclude that $\ker(\psi) = H/N$. From Theorem VIII.2.1 we now deduce

$$(G/N)/(H/N) = (G/N)/\ker(\psi) \cong \psi(G/N) = G/H,$$

which is what we wanted to prove. ∎

VIII.2.8 Example. The residue classes $2a \bmod 6$ for $a \in \mathbb{Z}$ form a normal subgroup in $\mathbb{Z}/6\mathbb{Z}$. This is precisely $2\mathbb{Z}/6\mathbb{Z}$, and Theorem VIII.2.7 applied to $G = \mathbb{Z}$ and $N = 6\mathbb{Z}$ and $H = 2\mathbb{Z}$ says that $(\mathbb{Z}/6\mathbb{Z})/(2\mathbb{Z}/6\mathbb{Z}) \cong \mathbb{Z}/2\mathbb{Z}$. ▬

VIII.3 Exercises

1. Determine all homomorphisms from $\mathbb{Z}/4\mathbb{Z}$ to $\mathbb{Z}/6\mathbb{Z}$.
2. Prove that $\mathbb{C}/\mathbb{Z}$ is isomorphic to the multiplicative group $(\mathbb{C}\setminus\{0\},\cdot,1)$.
3. Show that for $n,h\in\mathbb{Z}$ with $h\neq 0$ we have $h\mathbb{Z}/hn\mathbb{Z}\cong\mathbb{Z}/n\mathbb{Z}$.
4. With $N=\{(1),(1\,2)(3\,4),(1\,3)(2\,4),(1\,4)(2\,3)\}\subset S_4$, show $(S_4/N)/(A_4/N)\cong\mathbb{Z}/2\mathbb{Z}$.
5. Show that if $k|N$ then $(\mathbb{Z}/N\mathbb{Z})/(k\mathbb{Z}/N\mathbb{Z})\cong\mathbb{Z}/k\mathbb{Z}$. Is the condition $k|N$ necessary?
6. Let $n,m\in\mathbb{Z}$ be both positive. In the group D_{nm} we denote the counter clockwise rotation over $2\pi/nm$ by ρ. Take $\sigma=$ 'reflection in the x-axis' as element in D_{nm}, so $\sigma\rho\sigma=\rho^{-1}$. Consider $H=\{\mathrm{id},\sigma\}$ and $N=\{\mathrm{id},\rho^m,\rho^{2m},\ldots,\rho^{(n-1)m}\}$.
   (a) Show that H,N are subgroups of D_{nm} and N is normal in D_{nm}.
   (b) Prove that $HN\cong D_n$.
   (c) Prove that $D_m\cong D_{nm}/N$.
7. Let G be a group and $H_1,H_2\subset G$ normal subgroups. Define $\psi: G\to G/H_1\times G/H_2$ by $\psi(g)=(gH_1,gH_2)$.
   (a) Show that ψ is a homomorphism and $H_1\cap H_2$ is a normal subgroup in G.
   (b) Prove that $G/(H_1\cap H_2)$ is isomorphic to a subgroup of $G/H_1\times G/H_2$.
   (c) Use (b) to obtain a new proof of the Chinese Remainder Theorem.
8. We will show that for $n\geq 5$ the only normal subgroup in S_n different from $\{(1)\}$ or all of S_n, is A_n. Let N be such a nontrivial normal subgroup in S_n.
   (a) Use that A_n is simple and show that $A_n\subset N$ or $N\cap A_n=\{(1)\}$.
   (b) Show that in case $A_n\subset N$ it follows that $N=A_n$.
   (c) Show that if $N\neq\{(1)\}$ and $N\cap A_n=\{(1)\}$, then $NA_n=S_n$ and $S_n/N\cong A_n$.
   (d) Conclude in the situation of (c) that $\#N=2$, and prove this contradicts the assumption that $N\subset S_n$ is normal.

IX FINITELY GENERATED ABELIAN GROUPS

In this chapter we mainly consider commutative groups. Our main goal will be to present a description of all so-called finitely generated commutative groups.

IX.1 finitely generated groups

IX.1.1 Definition. A group G is called *finitely generated*, if elements $g_1,\ldots,g_n \in G$ exist with the property that every $g \in G$ can be written as

$$g = g_{i_1}^{\pm 1} \cdot \ldots \cdot g_{i_t}^{\pm 1}$$

with indices i_j such that $1 \le i_j \le n$ (note that it is allowed here that $i_k = i_\ell$, in other words any g_i can be used multiple times).

IX.1.2 Example.

1. Every finite group G is finitely generated, since in this case as $\{g_1,\ldots,g_n\}$ one may use the set of all elements in G.
2. The group $\mathbb{Z}^r = \mathbb{Z} \times \ldots \times \mathbb{Z}$ (the product of r copies of $\mathbb{Z}$) is finitely generated. Namely, take $e_1 = (1,0,\ldots,0)$, $e_2 = (0,1,0,\ldots,0),\ldots\ldots,e_r = (0,\ldots,0,1) \in \mathbb{Z}^r$. An arbitrary $(m_1,\ldots,m_r) \in \mathbb{Z}^r$ can be written as $m_1e_1 + \ldots + m_re_r$. (Moreover the integers m_j here are uniquely determined; the group $\mathbb{Z}^r$ is therefore an example of a so-called *free abelian group*, with basis $e_1,\ldots,e_r$.)
3. The additive group $(\mathbb{Q},+,0)$ is *not* finitely generated. Namely, given arbitrary $g_1,\ldots,g_n \in \mathbb{Q}$, the finite sum $\pm g_{i_1} \pm \ldots \pm g_{i_t}$ can be written as a/b with $a \in \mathbb{Z}$ and b equal to the least common multiple of the denominators of $g_1,\ldots,g_n$. Hence a number $c/d \in \mathbb{Q}$ with $c,d \in \mathbb{Z}$ and $\gcd(c,d) = 1$ and d larger than this least common multiple can not be expressed as a sum of $\pm g_i$'s. As a consequence, no finite set $\{g_1,\ldots,g_n\} \subset \mathbb{Q}$ generates all of $\mathbb{Q}$.
4. It is possible that a (nonabelian) group G is finitely generated whereas some subgroup $H \subset G$ is not. A nice example of this phenomenon is described in the paper B.L. van der Waerden, *Example d'un groupe avec deux générateurs, contenant un sous-groupe commutatif sans système fini de générateurs,* which appeared in the journal Nieuw Archief voor Wiskunde, Vol. **23** (1951), p. 190.
   In a similar (and in fact much easier) way it is possible that a finitely generated group is generated by elements of finite order, and yet the group contains elements of infinite order. As an example we take σ_1 the reflection in the x-axis and σ_2 = reflection in the line with equation $y = ax$. Then $\sigma_2\sigma_1$ is rotation over an angle α with $\tan(\alpha/2) = a$. For suitable a this rotation has infinite order. Clearly both σ_1 and σ_2 have order 2.

5. The group $\mathrm{SL}_2(\mathbb{Z})$ is finitely generated. As generators one can take the matrices $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. To show that indeed these matrices generate the group, consider an arbitrary $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$. Note that $T^q = \left(\begin{smallmatrix} 1 & q \\ 0 & 1 \end{smallmatrix}\right)$ and $S^2 = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. If $c = 0$ then $1 = \det(A) = ad$, hence $a = d = 1$ (and $A = T^b$) or $a = d = -1$ (and $A = S^2T^{-b}$). From now on we assume $c \neq 0$. Since $T^qA = \left(\begin{smallmatrix} a+qc & * \\ c & * \end{smallmatrix}\right)$, a suitable choice of q yields in the top left corner the remainder r of a upon division by c. For this choice of q one finds $ST^qA = \left(\begin{smallmatrix} -c & * \\ r & * \end{smallmatrix}\right)$. Repeating this process of multiplying by a suitable power of T and then by S makes the absolute value of the integers in the first column smaller and smaller. After finitely many steps one obtains a matrix $\left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right)$. As we saw earlier, this can be written as a combination of powers of S and T. it follows that also A can be expressed in this way.

∎

In the remainder of this chapter we restrict ourselves to commutative groups. The group law will be denoted by $+$.

IX.1.3 Theorem. *Any finitely generated commutative group $(A, +, 0)$ is isomorphic to a factor group $\mathbb{Z}^n/H$ for some subgroup $H \subset \mathbb{Z}^n$.*

Proof. Let the set $\{a_1, \ldots, a_n\}$ generate the group A. Define

$$\varphi \; : \; \mathbb{Z}^n \longrightarrow A$$

by $\varphi(m_1, \ldots, m_n) = m_1a_1 + \ldots + m_na_n$. It is easy to verify that φ is a homomorphism. Moreover φ is surjective because $a_1, \ldots, a_n$ generate A, so every element of A is an additive combination of the elements $\pm a_i$. Since A is commutative, here the order of the sequence of $\pm a_i$'s is irrelevant. So any $a \in A$ can be written as $a = n_1a_1 + \ldots + n_na_n$ which is the image of $(n_1, \ldots, n_n)$ under φ.

Put $H = \ker(\varphi)$. This is a subgroup of $\mathbb{Z}^n$. Theorem VIII.2.1 therefore states

$$\mathbb{Z}^n/H = \mathbb{Z}^n/\ker(\varphi) \cong \varphi(\mathbb{Z}^n) = A.$$

∎

IX.2 subgroups of free abelian groups

Theorem IX.1.3 shows that describing all finitely generated commutative groups boils down to describing all subgroups $H \subset \mathbb{Z}^n$ with the corresponding factor groups $\mathbb{Z}^n/H$. The case $n = 1$ we saw earlier (Example III.2.6): here $H = m\mathbb{Z}$ for some $m \geq 0$. So $\mathbb{Z}/H \cong \mathbb{Z}$ in case $m = 0$ and $\mathbb{Z}/H = (0)$ if $m = 1$, and $\mathbb{Z}/H = \mathbb{Z}/m\mathbb{Z}$ in general.

IX.2.1 Theorem. *If $H \subset \mathbb{Z}^n$ is a subgroup then $H \cong \mathbb{Z}^k$ for some k with $0 \leq k \leq n$.*

Proof. We use mathematical induction w.r.t. n. The case $n = 0$ is trivial, and the case $n = 1$ follows from Example III.2.6: namely, here $H = m\mathbb{Z}$ with $m \geq 0$. For $m = 0$ therefore $H = (0) \cong \mathbb{Z}^0$, and for $m > 0$ we have $\mathbb{Z} \cong m\mathbb{Z}$, with as explicit isomorphism multiplication by m.

As induction hypothesis, assume the theorem holds for $n \geq 1$. Let $H \subset \mathbb{Z}^{n+1}$ be a subgroup. Define

$$\pi \; : \; \mathbb{Z}^{n+1} \longrightarrow \mathbb{Z} \text{ by } \; \pi(m_1, \ldots, m_{n+1}) = m_{n+1}.$$

This is a homomorphism with kernel all sequences $(m_1,\ldots,m_{n+1}) \in \mathbb{Z}^{n+1}$ such that $m_{n+1} = 0$. We can therefore identify this kernel with $\mathbb{Z}^n$. Since $H \subset \mathbb{Z}^{n+1}$ is a subgroup, so is $H \cap \ker(\pi) \subset \mathbb{Z}^n$. Hence the induction hypothesis implies $H \cap \ker(\pi) \cong \mathbb{Z}^k$ for some k with $0 \le k \le n$.

Since $H \subset \mathbb{Z}^{n+1}$ is a subgroup, so is $\pi(H) \subset \mathbb{Z}$. Hence $\pi(H) = m\mathbb{Z}$ for some $m \ge 0$. If $m = 0$, then $\pi(H) = (0)$ so $H \subset \ker(\pi)$ which implies $\mathbb{Z}^k \cong H \cap \ker(\pi) = H$. So in this case the proof is complete. From now on we assume $m \neq 0$. Since $m \in m\mathbb{Z} = \pi(H)$, we have $h_{k+1} \in H$ with $\pi(h_{k+1}) = m$. Take $h_1,\ldots,h_k \in H \cap \ker(\pi)$ the images of $(1,0,\ldots,0),\ldots\ldots,(0,\ldots,0,1)$ under a chosen isomorphism $\mathbb{Z}^k \cong H \cap \ker(\pi)$. We will show that

$$\psi : \mathbb{Z}^{k+1} \longrightarrow H \text{ defined by } \psi(m_1,\ldots,m_{k+1}) = m_1h_1 + \ldots + m_{k+1}h_{k+1}$$

is an isomorphism. Evidently ψ is a homomorphism and $\mathbb{Z}^{k+1}$ is mapped by ψ into H. We first show that ψ is surjective. Take an arbitrary $h \in H$. Since $\pi(h) \in \pi(H) = m\mathbb{Z}$, we have $\pi(h) = m\ell$ for some $\ell \in \mathbb{Z}$. From this one deduces $\pi(h - \ell h_{k+1}) = \pi(h) - \pi(\ell h_{k+1}) = m\ell - \ell m = 0$, so $h - \ell h_{k+1} \in \ker(\pi) \cap H$. Using our isomorphism $\mathbb{Z}^k \cong \ker(\pi) \cap H$ we find $\ell_1,\ldots,\ell_k \in \mathbb{Z}$ with $h - \ell h_{k+1} = \ell_1 h_1 + \ldots + \ell_k h_k$. This shows $h = \psi(\ell_1,\ldots,\ell_k,\ell)$, so ψ is surjective.

Next we show that ψ is injective. By Theorem III.3.6 it suffices to verify that $\ker(\psi) = \{(0,\ldots,0)\} \subset \mathbb{Z}^{k+1}$. Let $(m_1,\ldots,m_{k+1}) \in \ker(\psi)$. Then $m_1h_1 + \ldots + m_{k+1}h_{k+1} = 0 \in H$, so $m_1h_1 + \ldots + m_kh_k = -m_{k+1}h_{k+1}$. From $h_1,\ldots,h_k \in \ker(\pi)$ it follows that $-m_{k+1}h_{k+1} \in \ker(\pi)$. Hence $0 = \pi(-m_{k+1}h_{k+1}) = -m_{k+1}m$. Our assumption $m \neq 0$ yields $m_{k+1} = 0$, so $m_1h_1 + \ldots + m_kh_k = -m_{k+1}h_{k+1} = 0$. Using $\mathbb{Z}^k \cong H \cap \ker(\pi)$ we conclude $(m_1,\ldots,m_k) = (0,\ldots,0) \in \mathbb{Z}^k$. So $m_1 = \ldots = m_k = m_{k+1} = 0$, which shows ψ is injective.

So ψ is an isomorphism, which finishes the induction argument. ∎

IX.2.2 Remark. The argument above repeatedly uses the fact that for a commutative group H one has $H \cong \mathbb{Z}^k$ if and only if $h_1,\ldots,h_k \in H$ exist such that every $h \in H$ can be written in a *unique* way as $h = m_1h_1 + \ldots + m_kh_k$. A group H having this property is called a free abelian group (with basis $h_1,\ldots,h_k$).

IX.2.3 Example. Consider $H \subset \mathbb{Z}^3$ given by

$$H = \{(a,b,c) \in \mathbb{Z}^3 \mid a + 2b + 3c \equiv 0 \bmod 6\}.$$

It is not hard to verify that H is a subgroup of $\mathbb{Z}^3$ (for example: H is the kernel of the homomorphism $\mathbb{Z}^3 \to \mathbb{Z}/6\mathbb{Z}$ given by $(a,b,c) \mapsto a+2b+3c \bmod 6$). By Theorem IX.2.1 and Remark IX.2.2 there exist $r \in \{0,1,2,3\}$ and $h_1,\ldots,h_r \in H$ such that H is the free abelian group with basis $h_1,\ldots,h_r$. We now determine such $r,h_1,\ldots,h_r$ by the method used in the proof of Theorem IX.2.1.

Let $\pi_i : \mathbb{Z}^3 \to \mathbb{Z}$ be the projection on the ith coordinate. We have $\pi_3(H) = \mathbb{Z}$ since $(1,1,1) \in H$ and hence $1 \in \pi_3(H)$; a subgroup of $\mathbb{Z}$ containing 1 equals $\mathbb{Z}$. The proof of Theorem IX.2.1 now shows that $(1,1,1)$ together with a basis for $\ker(\pi_3) \cap H$ yields a basis of H. By definition

$$\ker(\pi_3) \cap H = \{(a,b,0) \mid a + 2b \equiv 0 \bmod 6\}.$$

We find $\pi_2(\ker(\pi_3) \cap H) = \mathbb{Z}$, because $(4,1,0) \in \ker(\pi_3) \cap H$ and $\pi_2(4,1,0) = 1$. So a basis for $\ker(\pi_3) \cap H$ consists of $(4,1,0)$ together with a basis for $\ker(\pi_2) \cap \ker(\pi_3) \cap H$. We have

$$\ker(\pi_2) \cap \ker(\pi_3) \cap H = \{(a,0,0) \mid a \equiv 0 \bmod 6\} = \mathbb{Z}(6,0,0),$$

hence

$$H = \mathbb{Z} \cdot (6,0,0) + \mathbb{Z} \cdot (4,1,0) + \mathbb{Z} \cdot (1,1,1).$$

∎

We will now show that given any subgroup $H \subset \mathbb{Z}^n$, there is only one integer k with $H \cong \mathbb{Z}^k$, and moreover $0 \leq k \leq n$. This follows directly from Theorem IX.2.1 and the next result.

IX.2.4 Theorem. *If $\mathbb{Z}^k \cong \mathbb{Z}^\ell$, then $k = \ell$.*

Proof. Consider the composition $\mathbb{Z}^k \cong \mathbb{Z}^\ell \to \mathbb{Z}^\ell/2\mathbb{Z}^\ell$. Here the second map is the canonical homomorphism to a factor group, and $2\mathbb{Z}^\ell = 2\mathbb{Z} \times \ldots \times 2\mathbb{Z}$. This composition is a surjective homomorphism, and its kernel is $2\mathbb{Z}^k$. Hence Theorem VIII.2.1 implies $\mathbb{Z}^k/2\mathbb{Z}^k \cong \mathbb{Z}^\ell/2\mathbb{Z}^\ell$.

For any $m \geq 0$ one finds $\mathbb{Z}^m/2\mathbb{Z}^m \cong (\mathbb{Z}/2\mathbb{Z})^m$, since the homomorphism $\mathbb{Z}^m \to (\mathbb{Z}/2\mathbb{Z})^m$ given by $(n_1, \ldots, n_m) \mapsto (n_1 \bmod 2, \ldots, n_m \bmod 2)$ is surjective and has kernel $2\mathbb{Z}^m$; now apply Theorem VIII.2.1.

In our situation, combining the above arguments we find $(\mathbb{Z}/2\mathbb{Z})^k \cong (\mathbb{Z}/2\mathbb{Z})^\ell$. These groups have 2^k and 2^ℓ elements, respectively and therefore $k = \ell$. ∎

IX.2.5 Corollary. *If $H \subset \mathbb{Z}^n$ is a subgroup then a unique integer k exists with $H \cong \mathbb{Z}^k$ (and this k satisfies $0 \leq k \leq n$).*

Proof. By Theorem IX.2.1 k exists. If both k_1 and k_2 have the desired property then $\mathbb{Z}^{k_1} \cong H \cong \mathbb{Z}^{k_2}$, hence by Theorem IX.2.4 it follows that $k_1 = k_2$. ∎

IX.3 the structure of finitely generated abelian groups

The main theorem concerning finitely generated commutative groups is as follows.

IX.3.1 Theorem. *For any finitely generated commutative group there exist a unique integer $r \geq 0$ and a unique (possibly empty) finite sequence $(d_1, \ldots, d_m)$ with all $d_i \in \mathbb{Z}$ and $d_i > 1$ and $d_m | d_{m-1} | \ldots | d_1$, such that*

$$A \cong \mathbb{Z}^r \times \mathbb{Z}/d_1\mathbb{Z} \times \ldots \times \mathbb{Z}/d_m\mathbb{Z}.$$

IX.3.2 Definition. Given a finitely generated commutative group A, the integer r mentioned in Theorem IX.3.1 is called the *rank* of A. The integers $d_1, \ldots, d_m$ are called the *elementary divisors* of A.

IX.3.3 Example.

1. Any subgroup $H \subset \mathbb{Z}^n$ is by Corollary IX.2.5 isomorphic to $\mathbb{Z}^k$ for a unique k. In particular H this implies that H is finitely generated, and Theorem IX.3.1 holds here with $\operatorname{rank}(H) = k$ and an empty sequence of elementary divisors.
2. The finite commutative group $A = \mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z}$ clearly has $\operatorname{rank}(A) = 0$ and elementary divisors $(d_1, d_2) = (12, 2)$. Namely applying the Chinese Remainder Theorem (see Example III.3.2 4)

$$\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z} \cong \mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \cong \mathbb{Z}/12\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}.$$

—∎

The most important ingredient in the proof of Theorem IX.3.1 reads as follows.

IX.3.4 Theorem. *Given a subgroup $H \subset \mathbb{Z}^n$ with $H \neq (0)$, there exists a basis $f_1, \ldots, f_n$ for $\mathbb{Z}^n$ and an integer k with $1 \leq k \leq n$ and a sequence of integers $(d_1, \ldots, d_k)$ with $d_i > 0$ and $d_k | d_{k-1} | \ldots | d_1$ such that $d_1 f_1, \ldots, d_k f_k$ is a basis of H.*

Proof. Take a basis $e_1,\ldots,e_n$ for $\mathbb{Z}^n$ (say, the standard one) and a basis $g_1,\ldots,g_k$ for H (it exists by Theorem IX.2.1). Then $g_i = a_{1i}e_1 + \ldots + a_{ni}e_n$ $(i = 1,\ldots,k)$ for certain $a_{ij} \in \mathbb{Z}$. The integers a_{ij} form a matrix

$$\begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nk} \end{pmatrix}.$$

Each pair of bases $(\beta = \{e_1,\ldots,e_n\}, \gamma = \{g_1,\ldots,g_k\})$ yields in this way an $n \times k$ matrix with integer coefficienten, expressing how the basis γ is given in terms of the basis β. Replacing the basis β or the basis γ by a different one a different matrix is obtained. Our aim is to change these bases into a β' for $\mathbb{Z}^n$ and a γ' for H such that the resulting matrix is

$$\begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_k \\ 0 & \ldots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

for integers $(d_1,\ldots,d_k)$ with $d_i > 0$ and $d_k | d_{k-1} | \ldots | d_1$.

The next algorithm brings in finitely many steps our initial matrix into a matrix of the desired form. After presenting the algorithm, we will show that indeed it corresponds to a change of the two bases.

Algorithm.

step 1 If A is the zero matrix, we are done. If not, take (i,j) such that $|a_{ij}| > 0$ is minimal. Interchange the first and the i-th row as well as the first and the j-th column. The new matrix has $a_{11} \neq 0$ and in absolute value minimal. We now try to make the remaining entries in the first column zero.

step 2 If an integer a_{i1} in the first column (for $i \neq 1$) is nonzero, add a suitable multiple of the first row to the i-th row such that a_{i1} is replaced by an integer r with $0 \leq r < |a_{11}|$. Is $r \neq 0$, then interchange the i-th and the first row. Repeat this step until the first column only has a nonzero entry in place $(1,1)$.

step 3 Analogously, make the remaining entries in the first row equal to zero.

step 4 We now make sure that all entries in the matrix are multiples of a_{11}, as follows. If $a_{11} \nmid a_{ij}$, replace the i-th row by the sum of the i-th and the first (this only changes one entry in the first column), and add a suitable multiple of the first column to the j-th. This yields a_{ij} with $0 \leq a_{ij} < |a_{11}|$. It is $\neq 0$ since otherwise a_{ij} would have been divisible by a_{11}. Now start all over at step 1 with the new matrix. The new a_{11} obtained in step 1 is in absolute value strictly smaller than the old one, hence after finitely many steps indeed all a_{ij} are multiples of a_{11}.

step 5 Apply the steps 1 to 4 to the matrix obtained from the one found so far by deleting the first row and the first column. All entries of this smaller matrix are multiples of the a_{11} constructed above, and this property remains true during the steps. So at the end, the smaller matrix has in its top left corner and integer a_{22} which is a multiple of a_{11} and the remaining entries in its first row and column are 0. Moreover a_{22} divides a_{ij} for all $i,j \geq 3$. Continuing in this way results in a matrix with $a_{ij} = 0$ if $i \neq j$ and $a_{11} | a_{22} | \ldots | a_{kk}$.

step 6 Finally, multiply rows by ± 1 and put the first k vectors in the two bases in the reverse order to obtain a matrix as desired.

We now show that the changes made in the algorithm to the initial matrix correspond to changes of a basis for either $\mathbb{Z}^n$ or H. To this end, we describe some ways of changing a basis, and we explain the effect it has on the matrix.

1. Interchange in the basis for $\mathbb{Z}^n$ the j-th and the k-th basis vector.
   Obviously this results in a new basis for $\mathbb{Z}^n$. Een element $\sum_i a_i e_i \in \mathbb{Z}^n$ is in terms of the new basis given as $a_1e_1 + \ldots + a_k e_k + \ldots + a_j e_j + \ldots + a_n e_n$. Hence in the matrix this corresponds to interchanging the j-th and the k-th row.
2. Interchange in the basis for H the j-th and the k-th basis vector. The effect on the matrix is that the j-th and the k-th column are interchanged.
3. Replace the j-th basis vector e_j of $\mathbb{Z}^n$ by its opposite $-e_j$.
   This yields of course a new basis of $\mathbb{Z}^n$. With respect to the new basis an element of $\mathbb{Z}^n$ has as j-th coordinate -1 times the old j-th coordinate. Hence the effect on our matrix is that all entries in the j-th row are multiplied by -1.
4. Replace the j-th basis vector of the given basis for H by its opposite. The effect on the matrix is that all integers in the j-th column are multiplied by -1.
5. Let $a \in \mathbb{Z}$ and $i \neq j$ and replace in the basis for $\mathbb{Z}^n$ the basis vector e_i by $e_i' = e_i - ae_j$.
   Indeed this provides a new basis for $\mathbb{Z}^n$. Namely, the map $\mathbb{Z}^n \to \mathbb{Z}^n$ given by
   $$\sum_k a_k e_k \mapsto a_1e_1 + \ldots + a_i e_i + \ldots + (a_i a - a_j)e_j + \ldots + a_n e_n$$
   is an isomorphism of groups (check!), and this map sends $\{e_1, \ldots, e_i, \ldots, e_n\}$ to $\{e_1, \ldots, e_i', \ldots, e_n\}$.
   Since
   $$\sum_k a_k e_k = a_1e_1 + \ldots + a_i(e_i - ae_j) + \ldots + (a_j + a_i a)e_j + \ldots + a_n e_n$$
   it follows that the effect of this change of basis on the matrix is that the j-th row is replaced by the sum of th j-th and a times the i-th.
6. Finally, in the basis for H one can in an analogous way replace basis vector g_i by $g_i - ag_j$. Similar to the above, the effect on the matrix is that the j-th column is replaced by the j-th plus a times the i-th.

The conclusion from the base changes described here is that if the $n \times k$ matrix A expresses how a basis of H is represented with respect to a basis of $\mathbb{Z}^n$, and the matrix B is obtained from A by repeatedly executing the following steps:

1. interchange two rows or two columns in the given matrix;
2. multiply a row or a column by -1 in the given matrix;
3. add a times a *different* row/column to a given row/column in the given matrix;

then also B expresses how some basis for H is given in terms of some basis for $\mathbb{Z}^n$. This completes the proof of Theorem IX.3.4. ∎

IX.3.5 Remark. In fact the procedure presented above can be used in a more general context. Namely if $H \subset \mathbb{Z}^n$ is given as $H = \mathbb{Z}g_1 + \ldots + \mathbb{Z}g_k$ for certain $g_i \in \mathbb{Z}^n$, without assuming that the g_i are a basis for H, in exactly the same way a matrix can be made from the generators g_i. Applying the algorithm to this matrix changes it to a new one with only on the diagonal nonzero entries (unless $H = \{0\}$). The columns $\neq 0$ of the resulting matrix describe a basis for H, expressed in some basis for $\mathbb{Z}^n$. Hence this provides a way, given a finite set of generators for some subgroup of $\mathbb{Z}^n$, to obtain on the one hand a new proof of Theorem IX.2.1 for this subgroup and on the other hand to construct a basis for this subgroup.

Using Theorem IX.3.4 we will now show the existence of $r, d_1, \ldots, d_m$ as given in Theorem IX.3.1.

Proof. (existence of $r, d_1, \ldots, d_m$ in Theorem IX.3.1.) Let A be a finitely generated commutative group. By Theorem IX.1.3 $A \cong \mathbb{Z}^n/H$ for some subgroup $H \subset \mathbb{Z}^n$. Choose bases $f_1, \ldots, f_n$ of $\mathbb{Z}^n$ and $d_1 f_1, \ldots, d_k f_k$ of H as described in Theorem IX.3.4. Under the isomorphism $\mathbb{Z}^n \cong \mathbb{Z}^n$ given by $\sum a_i f_i \mapsto (a_1, \ldots, a_n)$ the subgroup H is mapped to the subgroup $d_1\mathbb{Z} \times d_2\mathbb{Z} \times \ldots \times d_k\mathbb{Z} \times (0) \times \ldots \times (0)$. Now consider

$$\varphi : \mathbb{Z}^n \longrightarrow \mathbb{Z}/d_1\mathbb{Z} \times \ldots \times \mathbb{Z}/d_k\mathbb{Z} \times \mathbb{Z}^{n-k}$$

given by $\varphi(a_1, \ldots, a_n) = (a_1 \bmod d_1, \ldots, a_k \bmod d_k, a_{k+1}, \ldots, a_n)$. This is a surjective homomorphism with kernel $d_1\mathbb{Z} \times d_2\mathbb{Z} \times \ldots \times d_k\mathbb{Z} \times (0) \times \ldots \times (0)$. Hence Theorem VIII.2.1 shows that

$$\begin{aligned} A \cong \mathbb{Z}^n/H &\cong \mathbb{Z}^n/(d_1\mathbb{Z} \times d_2\mathbb{Z} \times \ldots \times d_k\mathbb{Z} \times (0) \times \ldots \times (0)) \\ &\cong \mathbb{Z}/d_1\mathbb{Z} \times \ldots \times \mathbb{Z}/d_k\mathbb{Z} \times \mathbb{Z}^{n-k}. \end{aligned}$$

Removing the factors $\mathbb{Z}/1\mathbb{Z} \cong (0)$ from this product shows the existence stated in Theorem IX.3.1. ∎

It remains to verify the uniqueness of the integers $r, d_1, \ldots, d_m$ in Theorem IX.3.1. Here a useful concept is the following.

IX.3.6 Definition. Let A be an abelian group. The set $A_{\text{tor}} = \{a \in A \mid \text{ord}(a) < \infty\}$ is a subgroup of A called the *torsion subgroup* of A.

Verify yourself that indeed A_{tor} is a subgroup. It holds that the factor group A/A_{tor} contains, apart from its unit element, no elements of finite order (check!). Is A finitely generated then we know $A \cong \mathbb{Z}^r \times \mathbb{Z}/d_1\mathbb{Z} \times \ldots \times \mathbb{Z}/d_m\mathbb{Z}$ and here the second group has as elements of finite order precisely the elements $(0, \ldots, 0, \overline{a_1}, \ldots, \overline{a_m})$, for $\overline{a_i} \in \mathbb{Z}/d_i\mathbb{Z}$. As a consequence $A_{\text{tor}} \cong \mathbb{Z}/d_1\mathbb{Z} \times \ldots \times \mathbb{Z}/d_m\mathbb{Z}$. Moreover, $A/A_{\text{tor}} \cong \mathbb{Z}^r$: indeed, consider the composition

$$A \cong \mathbb{Z}^r \times \mathbb{Z}/d_1\mathbb{Z} \times \ldots \times \mathbb{Z}/d_m\mathbb{Z} \longrightarrow \mathbb{Z}^r$$

where the second map is projection on the first r coordinates. This is a surjective homomorphism, with kernel A_{tor}. Hence Theorem VIII.2.1 implies $A/A_{\text{tor}} \cong \mathbb{Z}^r$.

This discussion implies in particular that given a finitely generated abelian group A, the integer r in Theorem IX.3.1 equals the rank of the finitely generated free group A/A_{tor}. Hence by Theorem IX.2.4 it is unique.

We moreover conclude that $d_1 \cdot \ldots \cdot d_r = \#(\mathbb{Z}/d_1\mathbb{Z} \times \ldots \times \mathbb{Z}/d_m\mathbb{Z}) = \#A_{\text{tor}}$ and therefore the product of the integers d_1 up to d_m does not depend on the actual choice of integers as in Theorem IX.3.1.

Arguments like this will imply the uniqueness $d_1, \ldots, d_m$. We begin by showing some more properties of these integers. The number d_1 (this is the largest elementary divisor) is unique. Namely d_1 is the largest integer appearing as the order of an element in $\mathbb{Z}/d_1\mathbb{Z} \times \ldots \times \mathbb{Z}/d_m\mathbb{Z}$ (check!). This equals the maximal order of elements in A_{tor} which determines it.

Also the *number* of elementary divisors is fully determined by A. Namely, take a prime number p. Multiplying by p defines a homomorphism $A \to A$. Its kernel we write as $A[p]$. This is a subgroup of A and of A_{tor}. The number of elements in $A[p]$ equals the number of elements $(\overline{a_1}, \ldots, \overline{a_m}) \in \mathbb{Z}/d_1\mathbb{Z} \times \ldots \times \mathbb{Z}/d_m\mathbb{Z}$ with $p(\overline{a_1}, \ldots, \overline{a_m}) = (\overline{0}, \ldots, \overline{0})$. A small calculation shows that this equals p^k, with k the number of i's such that $p | d_i$. This k is maximal when $p \mid d_m$, in which case $k = m$. So we conclude that m equals the maximal exponent k such that a prime p exists with $\#A[p] = p^k$. This determines m in terms of A.

Proof. (of the uniqueness of the elementary divisors in Theorem IX.3.1.) We show the uniqueness of $d_1,\ldots,d_m$ by mathematical induction w.r.t. $\#A_{\text{tor}} = d_1\cdot\ldots\cdot d_m$. For $\#A_{\text{tor}} = 1$ the sequence of elementary divisors is empty so in this case uniqueness holds. Suppose $\#A_{\text{tor}} = N > 1$ and assume the uniqueness for all A' with $\#A'_{\text{tor}} < N$. Let $d_1,\ldots,d_m$ be a sequence of elementary divisors for A. We have $d_m > 1$ since $N > 1$. Take a prime $p|d_m$ and consider the factor group $A' = A/A[p]$. Since A is finitely generated, so is A' and $A'_{\text{tor}} \cong \mathbb{Z}/\frac{d_1}{p}\mathbb{Z}\times\ldots\times\mathbb{Z}/\frac{d_m}{p}\mathbb{Z}$. The induction hypothesis implies that the sequence $d_1/p,\ldots,d_m/p$ is unique, which implies the same is true for d_1 up to d_m. ∎

We finally discuss subgroups of $\mathbb{Z}^n$ generated by n elements. In particular we will decide when such a subgroup has finite index in $\mathbb{Z}^n$.

IX.3.7 Theorem. *Suppose that $H \subset \mathbb{Z}^n$ is a subgroup generated by n elements $g_1,\ldots,g_n$ and $g_i = a_{1i}e_1 + \ldots + a_{ni}e_n$ for some basis $\{e_1,\ldots,e_n\}$ of $\mathbb{Z}^n$. Let $A = (a_{ij})$ be the corresponding $n\times n$ matrix; then H has finite index in $\mathbb{Z}^n$ if and only if* $\det(A) \neq 0$. *If* $\det(A) \neq 0$ *holds then* $\#\mathbb{Z}^n/H = [\mathbb{Z}^n : H] = |\det(A)|$.

Proof. By the method described in the proof of Theorem IX.3.4 one transforms A into a diagonal matrix with nonnegative integers $d_1,\ldots,d_n$ on the diagonal. Note that the steps in this procedure can only change the sign of the determinant. In particular $|\det(A)| = d_1\cdot\ldots\cdot d_n$. The results of this Section show that $\mathbb{Z}^n/H \cong \mathbb{Z}/d_1\mathbb{Z}\times\ldots\times\mathbb{Z}/d_n\mathbb{Z}$. The latter group is finite precisely when all d_i's are different from 0. Is this the case then $[\mathbb{Z}^n : H] = \#\mathbb{Z}^n/H = d_1\cdot\ldots\cdot d_n = |\det(A)|$. ∎

IX.3.8 Example. Take $A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 2 & 2 \\ 3 & 4 & 2 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 2 & 2 \\ 3 & 4 & 2 \end{pmatrix}$.

One computes $\det(A) = 0$ and $\det(B) = 4$. For the subgroups $H_1 = A(\mathbb{Z}^3)$ and $H_2 = B(\mathbb{Z}^3)$ of $\mathbb{Z}^3$ we therefore conclude that the factor group $\mathbb{Z}^3/H_1$ is infinite and $\mathbb{Z}^3/H_2$ consists of 4 elements. The method described in the proof of Theorem IX.3.4 transforms A into a diagonal matrix with entries $1,2,0$ on the diagonal. Hence $\mathbb{Z}^3/H_1 \cong \mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$. Similarly B is transformed into the diagonal matrix with entries $1,2,2$. So $\mathbb{Z}^3/H_2 \cong \mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$.

The element of order 2 in $\mathbb{Z}^3/H_1$ is the class $(0,1,1)+H_1$. Namely the given element is not the zero element in $\mathbb{Z}^3/H_1$ because this would imply $(0,1,1)\in H_1$. Since the second coordinate of any element in H_1 is even, this is not the case. The order of $(0,1,1)+H_1$ is indeed 2 since $2\cdot((0,1,1)+H_1) = (0,2,2)+H_1 = (0,0,0)+H_1$, as $(0,2,2)\in H_1$.

Analogously, try to find the three distinct elements of order 2 in $\mathbb{Z}^3/H_2$! ∎

IX.4 Exercises

1. Show that the multiplicative group $(\mathbb{Q}\setminus\{0\},\cdot,1)$ is not finitely generated.
2. Show that if N is a normal subgroup of a finitely generated group G, then the factor group G/N is finitely generated as well.
3. Given $S=\left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)$ and $T=\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbb{Z})$. Compute $U=ST$ and show that $\mathrm{ord}(S)=4$ and $\mathrm{ord}(U)=6$ and S and U generate the group $\mathrm{SL}_2(\mathbb{Z})$.
4. Write the matrix $\left(\begin{smallmatrix}55&21\\34&13\end{smallmatrix}\right)$ as a product of powers of $S=\left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)$ and $T=\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$.
5. Show that $\mathrm{GL}_2(\mathbb{Z})$ is finitely generated and give explicit generators all having finite order (for instance three generators suffice, of order $2,4$, and 6, respectively).
6. Present an alternative proof for the fact that $\mathbb{Z}^{k_1}\ncong\mathbb{Z}^{k_2}$ in case $k_1\neq k_2$, by verifying (and using) that a basis for $\mathbb{Z}^k$ is in fact also a basis for the vector space $\mathbb{R}^k$ over $\mathbb{R}$.
7. Find a basis for the subgroup $H=\{(a,b,c,d)\mid a+b+c+d=0 \text{ en } a\equiv c \bmod 12\}$ of $\mathbb{Z}^4$.
8. Determine the rank and the elementary divisors of each of the following groups.
   (a) $\mathbb{Z}\times 17\mathbb{Z}\times\mathbb{Z}/18\mathbb{Z}\times\mathbb{Z}/12\mathbb{Z}$.
   (b) $(\mathbb{Z}/15\mathbb{Z})^*$
   (c) $(\mathbb{Z}/17\mathbb{Z})^*$
   (d) $\mathbb{Z}^3$ modulo the subgroup generated by $(1,2,0)$ and $(3,0,0)$.
   (e) A/H with $A\subset\mathbb{Z}^5$ the group of all 5-tuples with sum 0 and $H=A\cap B(\mathbb{Z}^5)$ where $B=\begin{pmatrix}-13&1&1&0&0\\1&-13&1&0&0\\1&1&-1&1&1\\0&0&1&-2&0\\0&0&1&0&-3\end{pmatrix}$.
9. Find the number of pairwise non-isomorphic commutative groups consisting of 72 elements.
10. (a) Use that $5^{2^{n+1}}-1=(5^{2^n}-1)(5^{2^n}+1)$ to prove that $5^{2^n}-1$ contains exactly $n+2$ factors 2.
    (b) Conclude from (a) that the order of $\overline{5}$ in $(\mathbb{Z}/2^n\mathbb{Z})^*$ equals 2^{n-2} (for $n\geq 2$).
    (c) Show that for $n\geq 2$ the map $a \bmod 2^n\mapsto a \bmod 4$ is a well defined surjective homomorphism from $(\mathbb{Z}/2^n\mathbb{Z})^*$ to $(\mathbb{Z}/4\mathbb{Z})^*$, and its kernel is the subgroup generated by $\overline{5}$.
    (d) Determine the number of elements of order ≤ 2 in $(\mathbb{Z}/2^n\mathbb{Z})^*$, and use this to find the rank and the elementary divisors of $(\mathbb{Z}/2^n\mathbb{Z})^*$.
11. (a) Prove that if A is a finite commutative group and p is a prime with $p\nmid \#A$, then $A/pA\cong(0)$.
    (b) Show that for $A=\mathbb{Z}/N\mathbb{Z}$ and p a prime with $p|N$ one has $A/pA\cong\mathbb{Z}/p\mathbb{Z}$.
    (c) Prove that if A is a finitely generated abelian group and p is a prime, then $\#A/pA=p^k$ where k equals the sum of the rank of A and the number of i's such that the elementary divisor d_i of A is divisible by p.
12. Let $d\geq 3$ be an integer. In this problem we study the polynomial X^2+X+d. Let $\alpha_d\in\mathbb{C}$ be a zero of this polynomial and define $A_d=\{a+b\alpha_d\mid a,b\in\mathbb{Z}\}\subset\mathbb{C}$.
    (a) Show that A_d is a subgroup of the additive group $(\mathbb{C},+,0)$ and $A_d\cong\mathbb{Z}^2$.
    (b) Take $\beta=a+b\alpha_d\in A_d$. Show that $\beta A_d=\{\beta\cdot\gamma\mid\gamma\in A_d\}$ is a subgroup of A_d, and if $\beta\neq 0$ then $\#A_d/\beta A_d=a^2-ab+db^2$.
    (c) Let a be an integer satisfying $0\leq a\leq d-2$ and a^2+a+d is not prime. Let p be the least of all primes dividing a^2+a+d. Prove that $p\leq d-1$.

(d) Given a and p as above, let $H = pA_d + (a-\alpha_d)A_d = \{p\gamma + (a-\alpha_d)\delta \mid \gamma, \delta \in A_d\}$. Show that $H \subset A_d$ is a subgroup generated by p and $p\alpha_d, a-\alpha_d, d+(a+1)\alpha$. Conclude that $\#A_d/H = p$.

(e) Use (b) and (d) to conclude that H is not of the form βA_d, for any $\beta \in A_d$. Show that H is closed under multiplication by elements of A_d; this means: for every $h \in H$ and every $\gamma \in A_d$ the product $h\gamma$ is an element of H.

(f) Conclude that if an integer a exists with $0 \le a \le d-2$ and $a^2 + a + d$ *not* a prime number, then A_d contains a subgroup which is closed under multiplication by elements of A_d, and this subgroup cannot be written as βA_d for some $\beta \in A_d$.

(g) It is a fact from "*algebraic number theory*" that A_{41} and A_{17} have the property that all subgroups closed under multiplication by all elements of A_d have the form βA_d. Draw a conclusion from this concerning the polynomials $X^2 + X + 17$ and $X^2 + X + 41$.

X APPENDIX

X.1 Symmetriegroups van de platonische lichamen.

We start by geometrically describing the orthogonal maps of $\mathbb{R}^3$.

X.1.1 Theorem. Is φ een orthogonale afbeelding op $\mathbb{R}^3$ met $\det(\varphi) = \epsilon$, dan geldt $\epsilon = \pm 1$.

Verder zijn er een lijn L door de oorsprong, en een vlak V door de oorsprong loodrecht op L, waarvoor geldt dat φ zowel L als V naar zichzelf afbeeldt.

φ werkt op V als een rotatie, en op L als vermenigvuldiging met ϵ.

Meetkundig gezien zegt dit resultaat heel precies hoe een orthogonale afbeelding op $\mathbb{R}^3$ eruitziet: is de determinant 1, dan is het een draaiing 'om een lijn L'. Is de determinant -1, dan draaien we niet alleen om een lijn L, maar bovendien spiegelen we in het vlak V loodrecht op L. Hoe we dat vlak V en die lijn L en de hoek waarover geroteerd wordt kunnen bepalen, zal blijken uit het bewijs dat nu gegeven wordt.

Proof. Ten aanzien van de standaardbasis voor $\mathbb{R}^3$ wordt de orthogonale afbeelding φ gegeven door een 3×3-matrix A. Deze matrix voldoet aan $A^*A = I$. Omdat $\det(A) = \det(A^*)$, volgt dan $\epsilon^2 = \det(A)^2 = \det(A^*A) = \det(I) = 1$ en dus $\epsilon = \pm 1$.

Het eigenwaardenpolynoom van A heeft graad 3, en dus heeft dit polynoom minstens één reëel nulpunt, dat we λ noemen. Dit is dan een eigenwaarde van φ bij een eigenvector v. Omdat φ afstandbehoudend is, geldt $||v|| = ||\varphi(v)|| = ||\lambda v|| = |\lambda| \cdot ||v||$, dus $\lambda = \pm 1$. Laat W het vlak door de oorsprong loodrecht op v zijn. We beweren dat φ dit vlak op zichzelf afbeeldt. Immers, laat $w \in W$ willekeurig. We moeten aantonen dat $\varphi(w) \in W$, hetgeen precies wil zeggen dat $\varphi(w) \perp v$. Welnu, $\varphi^*\varphi = \mathrm{id}$ en $\lambda = \pm 1$, en dus volgt door φ^* toe te passen op $v = \lambda\varphi(v)$ dat $\varphi^*(v) = \lambda v$. Dus $\langle \varphi(w), v \rangle = \langle w, \varphi^*(v) \rangle = \lambda \langle w, v \rangle = 0$, hetgeen we wilden laten zien.

De beperking van φ tot W is natuurlijk evenals φ zelf weer afstandbehoudend. Er zijn nu twee mogelijkheden voor deze beperking: het kan een rotatie zijn, en een spiegeling in een lijn in W door de oorsprong. In het geval van een rotatie zijn we direct klaar: ten aanzien van een basis van $\mathbb{R}^3$ bestaande uit de vector v samen met twee onderling loodrechte vectoren in W met lengte 1 wordt φ dan namelijk gegeven door een matrix $B = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$. Dan is $\epsilon = \det(A) = \det(B) = \lambda$ en we nemen $L =$ de lijn door v en de oorsprong, en $V = W$.

Is de beperking van φ tot W een spiegeling, neem dan w_1 een vector met lengte 1 op de lijn waarin gespiegeld wordt, en w_2 een vector in W met lengte 1 loodrecht

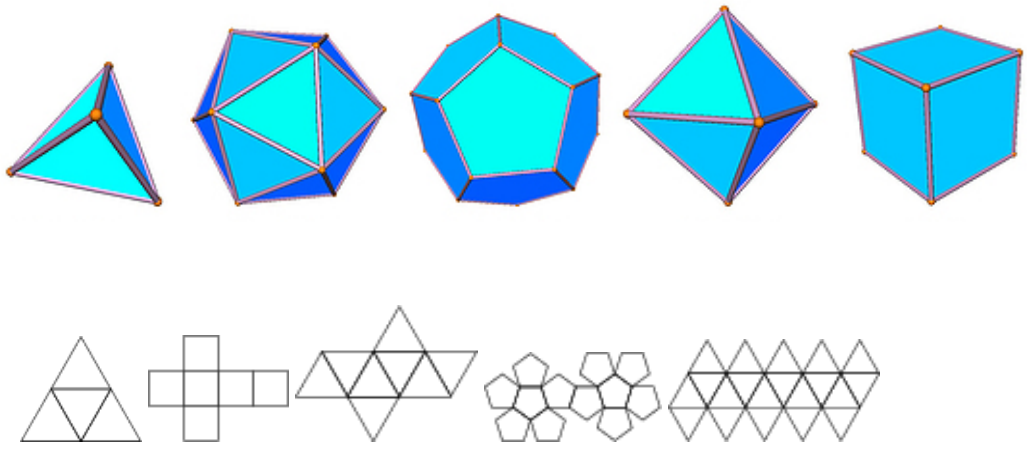
op w_1. Dan ziet ten aanzien van de basis v, w_1, w_2 voor $\mathbb{R}^3$ de matrix van φ eruit als $C = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$. Dus is $\epsilon = -\lambda$. Is $\lambda = 1$, dan kiezen we voor V het vlak door v en w_1, en voor L de lijn door w_2. (Dit correspondeert dus met de spiegeling in V.) Is $\lambda = -1$, dan wordt V het vlak door v en w_2, en L de lijn door w_1. (Dit levert dan draaien over 180 graden 'om de lijn L'.) ■

Precies zoals we dat hiervoor in het geval van een cirkel in $\mathbb{R}^2$ deden, beschouwen we nu een bol in $\mathbb{R}^3$. De symmetriegroep van de bol is de hele O(3) die zoals we zagen bestaat uit draaiingen en draaispiegelingen. Dit is een heel grote, niet-commutatieve groep.

Analoog aan de regelmatige n-hoeken in het vlak kan men vervolgens proberen regelmatige veelvlakken in de ruimte te maken. Dit gaat door op de bol n punten te kiezen zo, dat het figuur opgespannen door deze punten als zijvlakken steeds dezelfde regelmatige m-hoek heeft. We zullen een bewijs schetsen voor het feit dat dit slechts voor een paar combinaties (n, m) mogelijk is. Preciezer geformuleerd:

De enige ruimtelijke regelmatige n-vlakken zijn

- *De tetraeder, met* 4 *gelijkzijdige driehoeken als zijvlakken en met* 4 *hoekpunten en* 6 *ribben;*
- *De kubus, met* 6 *vierkanten als zijvlakken en met* 8 *hoekpunten en* 12 *ribben;*
- *De octaeder, met* 8 *gelijkzijdige driehoeken als zijvlakken en met* 6 *hoekpunten en* 8 *ribben;*
- *De dodecaeder, met* 12 *regelmatige vijfhoeken als zijvlakken en met* 20 *hoekpunten en* 30 *ribben;*
- *De icosaeder, met* 20 *gelijkzijdige driehoeken als zijvlakken en met* 12 *hoekpunten en* 30 *ribben.*



Schets van een bewijs. De symmetriegroep van een willekeurig regelmatig veelvlak is eindig. Immers, zo'n symmetrie wordt volledig bepaald door wat er met de hoekpunten van de figuur gebeurt (die hoekpunten corresponderen met vectoren die tezamen $\mathbb{R}^3$ opspannen). Dus er bestaan hooguit zoveel symmetrieën als er permutaties van de hoekpunten zijn, en dat aantal is eindig. Dit argument laat in feite zien dat de symmetriegroep van zo'n figuur isomorf is met een ondergroep van S_n, waarbij n het aantal hoekpunten is. Verder geldt, dat ieder hoekpunt door een geschikt element van de symmetriegroep op ieder willekeurig ander hoekpunt

kan worden afgebeeld. Dit kan zelfs al met een rotatie. Immers, stellen we ons de hoekpunten voor als punten op een bol, dan is door het draaien van die bol elk hoekpunt op de noordpool te leggen, en door de bol dan nog om de noord-zuid as te draaien 'behoudt het veelvlak z'n oorspronkelijke stand'. Ook merken we op, dat er bij een hoekpunt altijd rotaties zijn, die dat hoekpunt vasthouden. En vanwege Lemma V.2.6, toegepast op de rotaties van een vlak loodrecht op de lijn door de oorsprong en het gegeven hoekpunt, vormen die een groep isomorf met $\mathbb{Z}/m\mathbb{Z}$ voor zekere m. Deze draaiingen voeren de zijvlakken die in het gegeven hoekpunt bij elkaar komen cyklisch in elkaar over; in het bijzonder zijn er dus precies m zulke zijvlakken.

Bovenstaande observaties leiden tot de volgende strategie voor het bepalen van alle regelmatige veelvlakken. Bepaal eerst alle eindige ondergroups van de groep van draaiingen SO(3). Voor de gevonden groups bepalen we alle punten op een boloppervlak die door meerdere elementen van de groep worden vastgehouden. Dat levert alle mogelijke hoekpunten, en door uitgaande van één ervan alle beelden onder de groep te bepalen vinden we dan alle mogelijke regelmatige veelvlakken. We voeren dat nu in iets meer detail uit.

Laat G een willekeurige eindige ondergroep van SO(3) zijn bestaande uit precies $N \geq 2$ elementen. Met B duiden we de verzameling punten in $\mathbb{R}^3$ aan die op afstand 1 van de oorsprong liggen; dus B is het boloppervlak. Is $\sigma \in G$, en $\sigma \neq \mathrm{id}$, dan is σ vanwege Stelling X.1.1 een rotatie om een lijn L. Er zijn dus precies twee punten van B die door σ op zichzelf worden afgebeeld, namelijk de snijpunten van L met B. In totaal vinden we zo een verzameling van precies $2(N-1)$ paren (σ, P), waarbij $\sigma \in G, \sigma \neq \mathrm{id}, P \in B$ en $\sigma(P) = P$.

We bekijken vervolgens de zo verkregen punten $P \in B$. Bij zo'n P hebben we minstens één draaiing $\sigma \neq \mathrm{id}$ met $\sigma \in G$ en $\sigma(P) = P$. Alle rotaties in G die P vasthouden, houden ook de lijn door P en de oorsprong O vast. Bijgevolg kunnen we ze opvatten als eindige groep van rotaties van het vlak door de oorsprong loodrecht op de lijn door OP; vanwege Lemma V.2.6 is dit een groep G_P isomorf met $\mathbb{Z}/m_P\mathbb{Z}$, voor zekere $m_P \geq 2$. Omdat G_P een ondergroep is van G, geldt $m_P = \#G_P | \#G = N$. Schrijf $N = m_P \cdot n_P$. Uit het bewijs van Stelling III.2.7 concluderen we dat er $\sigma_1, \ldots, \sigma_{n_P} \in G$ zijn zodat $G = \sigma_1 G_P \cup \ldots \cup \sigma_{n_P} G_P$, waarbij geldt dat $\sigma_i G_P \cap \sigma_j G_P = \varnothing$ voor $i \neq j$. Ieder element van $\sigma_i G_P$ is te schrijven als $\sigma_i \tau$ voor een $\tau \in G_P$, en dus $\sigma_i \tau(P) = \sigma_i(P)$. Schrijf $P_i = \sigma_i(P)$. Er geldt $P_i \neq P_j$ als $i \neq j$. Immers, anders zou $\sigma_i(P) = \sigma_j(P)$ en dus $\sigma_i^{-1}\sigma_j \in G_P$, hetgeen impliceert dat $\sigma_i G_P$ en $\sigma_j G_P$ geen lege doorsnede zouden hebben. Dus P wordt door de elementen van G op in totaal n_P verschillende punten afgebeeld. De zo verkregen verzameling van n_P punten noteren we als C_P.

Neem P als boven, $\sigma \in G$ willekeurig, en $Q = \sigma(P)$. Dan geldt $m_P = m_Q$. Immers, is $\tau \in G_P$, dan is $\sigma\tau\sigma^{-1} \in G_Q$ en omgekeerd is voor $\rho \in G_Q$ het element $\sigma^{-1}\rho\sigma \in G_P$. Dit levert een bijectie (zelfs een isomorfisme van groups) tussen G_P en G_Q, dus in het bijzonder hebben beide hetzelfde aantal elementen oftewel $m_P = m_Q$. Dit levert een tweede manier om het aantal paren (σ, P) met $\sigma \in G, \sigma \neq \mathrm{id}, P \in B$ en $\sigma(P) = P$ te tellen: het is gelijk aan $\sum_C n_C(m_C - 1)$. Hier sommeren we over alle verschillende verzamelingen $C = C_P$, en voor zo'n C_P is n_C het aantal punten $Q \in C_P$ en m_C het aantal elementen in G_Q voor elke $Q \in C_P$.

De twee uitdrukkingen voor het aantal paren leveren de gelijkheid

$$2N - 2 = \sum_C (N - \frac{N}{m_C}).$$

Dit delen door N geeft $2 - \frac{2}{N} = \sum_C (1 - 1/m_C)$. De rest van het bewijs bestaat voornamelijk uit het analyseren van deze vergelijking. Het linkerlid is groter dan 1, dus moet de som in het rechterlid uit minstens twee termen bestaan. Omdat het linkerlid kleiner is dan 2, en het rechterlid bestaat uit termen die minstens 1/2 zijn, bestaat de som daar dus óf uit 2, óf uit drie termen. Eerst zullen we het geval

dat de som eruitziet als $2-2/N=(1-1/m_1)+(1-1/m_2)$ beschouwen. Met N vermenigvuldigen levert $2=N/m_1+N/m_2$. Omdat m_1 en m_2 positieve delers van N zijn, moet gelden $m_1=m_2=N$. Elk van de twee verzamelingen C_i bestaat in dit geval uit $n_i=N/m_i=1$ punt. Omdat de elementen van G rotaties zijn en allemaal deze twee punten vasthouden, is G een eindige groep van rotaties om een vaste lijn, dus met behulp van Lemma V.2.6 volgt dat $G\cong\mathbb{Z}/N\mathbb{Z}$. Ook concluderen we dat dit geval ons geen regelmatig veelvlak oplevert, want we zouden slechts één hoekpunt krijgen.

In het resterende geval zijn er drie verzamelingen C_i. We schrijven weer $m_i=m_{C_i}$ en ook $n_i=\#C_i=N/m_i$. De volgorde van de C_i's kiezen we zo, dat $m_1\le m_2\le m_3$. De vergelijking die we hebben kan geschreven worden als

$$1+\frac{2}{N}=\frac{1}{m_1}+\frac{1}{m_2}+\frac{1}{m_3}.$$

Zouden alle m_i minstens 3 zijn, dan was het rechterlid van deze gelijkheid ≤ 1 en dat kan niet. Dus geldt $m_1=2$. Met andere woorden,

$$\frac{1}{2}+\frac{2}{N}=\frac{1}{m_2}+\frac{1}{m_3}.$$

Hetzelfde argument als zojuist levert dat m_2,m_3 niet beide minstens 4 kunnen zijn. Dus is $m_2=2$ of $m_2=3$. Eerst analyseren we de mogelijkheid $m_2=2$. In dit geval geldt $N=2m_3$, en m_3 mag willekeurig ≥ 2 zijn, zeg $m_3=n$. De verzamelingen C_1,C_2 bestaan hier elk uit $N/m_1=n$ punten, en ieder van die punten wordt vastgehouden door een ondergroep van G bestaande uit $m_1=2$ rotaties. In zo'n ondergroep zit dus een draaiing over 180 graden om de lijn door de oorsprong en dat punt. Hieruit volgt dat bijvoorbeeld de n punten in C_1 alle in één vlak V door de oorsprong liggen. De doorsnede van V met B is een cirkel, en omdat G uit isometrieën bestaat, vormen de punten in C_1 dan een regelmatige n-hoek. De groep G beeldt V op zichzelf af en in het bijzonder deze n-hoek ook. 'Beperken tot V' is een homomorfisme van G naar de symmetriegroep van de n-hoek, en dat is D_n. Dit homomorfisme is injectief, want als voor een draaiing geldt dat z'n beperking tot een vlak de identiteit is, dan is de draaiing zelf de identiteit. Dus het beperkingshomomorfisme heeft als kern alleen de identiteit en is bijgevolg injectief. Daar zowel G als D_n uit $2n$ elementen bestaan, volgt $G\cong D_n$. Als ondergroep van SO(3) ziet D_n er volgens bovenstaande analyse als volgt uit. Kies een vlak V door de oorsprong en een regelmatige n-hoek om O in V. De rotaties in D_n zijn dan rotaties om de lijn door O loodrecht op V, en de spiegelingen zijn de draaiingen over 180 graden om een lijn door O en een hoekpunt van de n-hoek.

De resterende mogelijk is $m_1=2$ en $m_2=3$. In dat geval geldt $m_3\ge m_2=3$ en $1/6+2/N=1/m_3$. Dus kan m_3 niet meer dan 5 zijn. Er resteren drie mogelijkheden, namelijk $m_3=3$ en $m_3=4$ en $m_3=5$. Is $m_3=3$, dan geldt $N=12$. De verzamelingen C_2 en C_3 bestaan in dit geval elk uit 4 punten. Ieder van deze punten wordt door een groep van orde 3, dus door rotatie over 120 graden vastgehouden. Dit leidt tot een regelmatig figuur bestaande uit 4 hoekpunten, waar steeds 3 ribben en 3 zijvlakken samenkomen. Dit is precies de tetraeder.

Wanneer $m_3=4$, dan is $N=24$. De verzameling C_2 spant dan een figuur op met $24/3=8$ hoekpunten. In elk van deze punten komen 3 zijden en 3 ribben bij elkaar. Deze worden in elkaar overgevoerd door rotaties over veelvouden van 120 graden. Dit bepaalt weer precies een regelmatig veelvlak, namelijk de kubus. Bezien we de verzameling C_3, dan geeft dat $24/4=6$ punten waarin 4 ribben/zijvlakken samenkomen. C_3 spant dus een octaeder op.

Tenslotte $m_3=5$, wat leidt tot $N=60$. Het figuur opgespannen door C_2 heeft nu 20 hoekpunten. In ieder ervan komen 3 ribben/vlakken samen, en we krijgen een dodecaeder. Op dezelfde manier geeft C_3 hier aanleiding tot een icosaeder. Hiermee zijn de regelmatige veelvlakken geklassificeerd, en we hebben zelfs voor ieder

ervan het aantal elementen van de ondergroep van hun symmetriegroep bestaande uit alle rotaties bepaald. □

We gaan nu voor elk van de regelmatige veelvlakken een beschrijving van de symmetriegroep geven. Allereerst merken we op dat (met uitzondering van het geval van de tetraeder) de afbeelding -1 (puntspiegelen in de oorsprong) hier een element van is. Verder is het een eindige ondergroep van O(3). Is τ zo'n symmetrie, dan $\det(\tau) = \pm 1$. In het geval $\det(\tau) = 1$ is τ een rotatie. Is $\det(\tau) = -1$, dan is $-\tau$ een rotatie. Hieruit concluderen we dat de symmetriegroep precies bestaat uit alle rotaties τ, en alle afbeeldingen $-\tau$.

Voor de tetraeder bestaat eenzelfde soort redenering: neem een spiegeling σ in de symmetriegroep ervan. Dan geldt $\det(\sigma) = -1$ en $\sigma^2 = \mathrm{id}$. Is een τ in de symmetriegroep *geen* rotatie, dan $\det(\tau) = -1$ en $\tau = \sigma \cdot \sigma\tau$ waarbij $\sigma\tau$ *wel* een rotatie is, omdat immers $\det(\sigma\tau) = \det(\sigma)\det(\tau) = -1 \cdot -1 = 1$. Dus de hele symmetriegroep bestaat uit de rotaties, en σ maal deze rotaties.

In het bijzonder bestaat de symmetriegroep van een regelmatig veelvlak uit resp. $24, 48, 48, 120, 120$ elementen.

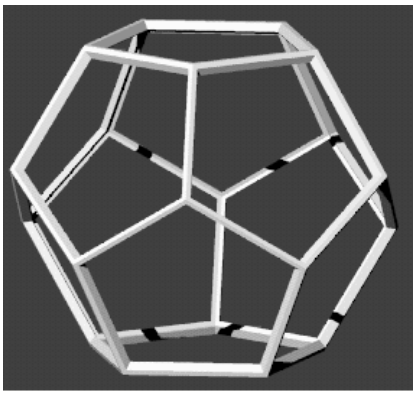
De tetraeder. De symmetriegroep is isomorf met S_4. Immers, gezien als permutatiegroep op de hoekpunten is het een ondergroep van S_4, en omdat de groep uit 24 elementen bestaat is het dan de hele S_4.

De kubus en de octaeder. Deze twee figuren hebben dezelfde symmetriegroep. Immers, leggen we binnen de kubus een bol die precies aan de zijvlakken raakt, dan vormen de 6 raakpunten precies de hoekpunten van een octaeder. Hieruit volgt dat een symmetrie van de kubus ook een symmetrie van de octaeder oplevert. Dit argument kan vervolgens omgedraaid worden: een ingeschreven bol aan de octaeder raakt deze weer precies in de hoekpunten van een kubus. Dus beide figuren hebben precies dezelfde symmetrieën.

Zoals we gezien hebben heeft deze symmetriegroep 48 elementen. De groep is isomorf met $S_4 \times \{\pm 1\}$. Dit kan als volgt bewezen worden. Een kubus heeft 4 hoofddiagonalen, en deze worden door de symmetriegroep gepermuteerd. Verder heeft elke symmetrie een determinant ± 1. Aldus wordt een homomorfisme naar $S_4 \times \{\pm 1\}$ verkregen. Zit τ in de kern van dit homomorfisme, dan houdt τ alle hoofddiagonalen vast en $\det(\tau) = 1$. τ is dan een rotatie om één van de hoofddiagonalen, en het is niet moeilijk om na te gaan dat het feit dat τ de drie andere diagonalen naar zichzelf moet sturen impliceert dat $\tau = \mathrm{id}$. Dus het gegeven homomorfisme is injectief, en omdat zowel de symmetriegroep als $S_4 \times \{\pm 1\}$ uit 48 elementen bestaan, zijn beide groups isomorf.

De dodecaeder en de icosaeder. Hetzelfde argument dat voor de kubus en de octaeder gegeven werd toont aan dat deze twee dezelfde symmetriegroep hebben. In dit geval is de symmetriegroep isomorf met $A_5 \times \{\pm 1\}$. We gebruiken de dodecaeder om dit aan te tonen.

Nummer de ribben van het bovenvlak $1, 2, 3, 4, 5$. Voor i met $1 \leq i \leq 5$ definiëren we V_i als de verzameling ribben van de dodecaeder die in een richting wijzen die óf evenwijdig is, óf loodrecht staat op de richting van ribbe i. Elke V_i bestaat dan uit precies 6 ribben. Omdat symmetrieën hoeken in even grote hoeken overvoeren, werkt de symmetriegroep van de dodecaeder als permutaties op de verzameling $\{V_1, V_2, V_3, V_4, V_5\}$. De symmetrie 'puntspiegelen in de oorsprong' houdt elk van de V_i's op z'n plaats, dus om na te gaan welke permutaties voorkomen als beeld van een symmetrie hoeven we alleen maar naar draaiingen te kijken. Men ziet dan eenvoudig in dat alleen *even* permutaties voorkomen. Door ook nog een symmetrie



te sturen naar z'n determinant krijgen we een homomorfisme naar $A_5 \times \{\pm 1\}$. De kern blijkt alleen uit de identiteit te bestaan, dus omdat zowel de symmetriegroep als $A_5 \times \{\pm 1\}$ precies 120 elementen heeft zijn ze isomorf.

X.2 Exercises

1. Deze opgave is bedoeld om de draaiingshoek die voorkomt in orthogonale afbeeldingen op $\mathbb{R}^3$ te bepalen. Zie het bewijs van Stelling X.1.1 voor de gebruikte notaties. De matrix A in dat bewijs vatten we hier op als complexe 3×3-matrix.
   (a) Laat zien dat de eigenwaarden van A van de vorm $\lambda = \pm 1, e^{i\alpha}$ en $e^{-i\alpha}$ zijn, met $0 \leq \alpha < 2\pi$.
   (b) Is $e^{i\alpha}$ een niet-reële eigenwaarde van A, bewijs dan dat een eigenvector hierbij te schrijven is als $x + iy$ voor reële vectoren x, y en dat x, y en de eigenvector bij λ een basis van $\mathbb{R}^3$ vormen van onderling loodrechte vectoren.
   (c) Ga na dat A de ruimte opgespannen door x, y naar zichzelf afbeeldt, en dat A op die ruimte werkt als een draaiing over de hoek α.
2. Lees de beschrijving van de symmetriegroep van de dodecaeder in dit dictaat. Kies vervolgens een hoekpunt P, en ga na welke permutaties in A_5 de rotaties om de lijn OP opleveren. Doe hetzelfde voor de rotaties om de lijn door een middelpunt van een zijvlak, en ook voor de rotatie (over 180 graden) om de lijn door het midden van een ribbe.
3. De tetraeder bevat precies 3 paren R_1, R_2, R_3 van elkaar niet snijdende ribben, en de symmetriegroep permuteert deze drie. Geef het homomorfisme: $S_4 \to S_3$ waartoe dit aanleiding geeft expliciet. Laat zien dat het een surjectie is, en beschrijf de kern.